

# HOW ROBUST ARE THE SIZE MEASUREMENTS OF HIGH-REDSHIFT COMPACT GALAXIES?

ROOZBEH DAVARI<sup>1,3</sup>, LUIS C. HO<sup>2,3</sup>, CHIEN Y. PENG<sup>4</sup>, AND SONG HUANG<sup>5</sup>

*Accepted for publication in The Astrophysical Journal*

## ABSTRACT

Massive quiescent galaxies at  $z \approx 2$  are apparently much more compact than galaxies of comparable mass today. How robust are these size measurements? We perform comprehensive simulations to determine possible biases and uncertainties in fitting single-component light distributions to real galaxies. In particular, we examine the robustness of the measurements of the luminosity, size, and other structural parameters. We devise simulations with increasing realism to systematically disentangle effects due to the technique (specifically using GALFIT) and the intrinsic structures of the galaxies. By accurately capturing the detailed substructures of nearby elliptical galaxies and then rescaling their sizes and signal-to-noise to mimic galaxies at different redshifts, we confirm that the massive quiescent galaxies at  $z \approx 2$  are significantly more compact intrinsically than their local counterparts. Their observed compactness is not a result of missing faint outer light due to systematic errors in modeling. In fact, we find that fitting multi-component galaxies with a single Sérsic profile, the procedure most commonly adopted in the literature, biases the inferred sizes higher by up to 10%–20%, which accentuates the amount of size evolution required. If the sky estimation has been done robustly and the model for the point-spread function is fairly accurate, GALFIT can retrieve the properties of single-component galaxies over a wide range of signal-to-noise ratios without introducing any systematic errors.

*Subject headings:* galaxies: elliptical and lenticular, cD — galaxies: formation — galaxies: photometry — galaxies: structure — galaxies: surveys

## 1. INTRODUCTION

Galaxy morphology has always been an important observable for understanding the formation and evolution of galaxies. Recent imaging studies using the *Hubble Space Telescope* (*HST*)/Wide Field Camera 3 (WFC3) and Advanced Camera for Surveys (ACS) have broadened our understanding of the formation and evolution of galaxies. The advent of WFC3 has given access to the rest-frame optical light of galaxies at  $z \approx 2$ . Morphological studies of these galaxies show that quiescent systems make up a considerable fraction of the massive galaxy population at  $z \approx 2$  (e.g., Franx et al. 2003; Daddi et al. 2005; Kriek et al. 2006). Their structural evolution has been the subject of considerable interest, focusing in particular on their extremely compact nature compared to low-redshift galaxies of similar mass (e.g., Daddi et al. 2005; Toft et al. 2007; Trujillo et al. 2007; Buitrago et al. 2008; Cimatti et al. 2008; Franx et al. 2008; van der Wel et al. 2008; van Dokkum et al. 2008; Damjanov et al. 2009; Hopkins et al. 2009; Cassata et al. 2010, 2011; Mancini et al. 2010; Newman et al. 2012; Szomoru et al. 2012). The sizes of these “red nuggets,” less than  $\sim 1$  kpc, seem comparable to and sometimes even smaller than the point-spread function (PSF). The early formation and subsequent size growth of these massive, compact objects present a chal-

lenge to current models of galaxy formation and evolution (e.g., Wuyts et al. 2010; Oser et al. 2012). It is not clear via what pathways they become the massive galaxies of today. The rarity of these massive, compact galaxies at low redshift implies considerable size evolution between  $z = 2$  and  $z = 0$  (van Dokkum et al. 2008; Trujillo et al. 2009; Taylor et al. 2010; but see Saracco et al. 2010; Valentinuzzi et al. 2010; Poggianti et al. 2013). On average, from  $z \approx 2$ , these objects would have to increase their size by 3–4 times while doubling their stellar mass (van Dokkum & Brammer 2010; see also Ichikawa et al. 2012). However, efforts to accurately quantify this evolution are hindered by potential uncertainties in measurement techniques: the mass densities may simply be systematically overestimated due to modeling uncertainties of photometric masses, and/or the sizes may be underestimated due to a lack of imaging depth or other measurement issues (Hopkins et al. 2009; Muzzin et al. 2009).

In addition to the basic question of how these high- $z$  galaxies evolve in size, there is also still much debate about how these systems evolve in terms of their fundamental morphological type. By performing GALFIT fitting on a sample of 14 compact, massive galaxies at  $z = 2$ , van der Wel et al. (2011) find that a significant subset of their sample appears highly flattened in projection, which, considering viewing angle statistics, implies those galaxies have pronounced disks. They claim that  $65\% \pm 15\%$  of the population of massive quiescent at  $z > 2$  galaxies are disk-dominated. Bruce et al. (2012) find that a considerable fraction ( $25\% \pm 6\%$  using a definition of bulge-to-total ratio  $B/T < 0.5$  and  $40\% \pm 7\%$  using a definition of  $n < 2.5$  for disk-dominated) of the most quiescent galaxies, with specific star formation

<sup>1</sup> University of California, Riverside 900 University Avenue, Riverside, CA 92521, USA

<sup>2</sup> Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, P. R. China

<sup>3</sup> The Observatories of the Carnegie Institution for Science 813 Santa Barbara Street, Pasadena, CA 91101, USA

<sup>4</sup> Giant Magellan Telescope Organization 251 South Lake Avenue, Suite 300 Pasadena, CA 91101, USA

<sup>5</sup> School of Space Science and Astronomy, Nanjing University, Nanjing 210093, P. R. China

rates  $< 10^{-10} \text{ yr}^{-1}$ , have disk-dominated morphologies, including a small number of essentially pure disk galaxies (with  $B/T < 0.1$ ). They claim that these passive disks appear to be normal disks in the sense that they show an axial-ratio distribution comparable to that displayed by present-day disks. This implies that while the massive galaxy population is undergoing dramatic changes at this crucial epoch, the physical mechanisms that quench star formation activity are not obviously connected to those responsible for transforming the morphologies of massive  $z \approx 2$  galaxies into present-day giant ellipticals.

A recent morphological study of nearby elliptical galaxies has opened a new window to understanding the formation and assembly of early-type galaxies. Huang et al. (2013a, hereafter H13) present a detailed, comprehensive structural analysis of about 100 representative, nearby elliptical galaxies spanning a range of environments and stellar masses ( $M_* \approx 10^{10.2} - 10^{12.4} M_\odot$ ). They use **GALFIT** 3.0 (Peng et al. 2002, 2010) to perform two-dimensional, multi-component decomposition of relatively deep, moderately high-resolution  $V$ -band images acquired as a part of the Carnegie-Irvine Galaxy Survey<sup>6</sup> (CGS; Ho et al. 2011). H13 challenge the conventional notion that the main body of giant ellipticals follows a single structure described by a high Sérsic (1968) index (e.g.,  $n \geq 3 - 4$ ). They propose that the global light distribution of the majority ( $> 75\%$ ) of ellipticals are best described by three Sérsic components: a compact, inner core with typical effective radius  $R_e < 1$  kpc and luminosity fraction  $f \approx 0.1 - 0.15$ ; an intermediate-scale, middle component with  $R_e \approx 2.5$  kpc and  $f \approx 0.2 - 0.25$ ; and an extended, outer envelope with  $R_e \approx 10$  kpc and  $f \approx 0.6$ . The subcomponents have relatively low Sérsic indices, in the range  $n \approx 1 - 2$ . They also find that the ellipticity of the isophotes systematically increases toward large radii. They believe that the combination of their model’s inner and middle components for high-luminosity ellipticals resembles the compact, massive galaxies at high- $z$  (Huang et al. 2013b).

Despite there being codes that can automate the measurement of galaxy properties (e.g., Simard et al. 2002, 2011; Barden et al. 2012), there are still many disagreements over the analysis techniques, because they are often viewed as being too simple, and what the measurements mean. Even stepping up the sophistication, however, makes interpretation non-trivial unless one has physical motivations for doing so (H13). These disagreements inspire other techniques (Conselice 1997; Lotz et al. 2004) to measure galaxies non-parametrically, although each one has its own benefits and shortcomings. In addition to galaxy shapes being difficult to quantify, many other factors complicate the image analysis of galaxies, including determination of the sky background, the brightness of the galaxy [signal-to-noise ratio ( $S/N$ )], the resolution of the images,  $(1+z)^4$  surface brightness dimming, finding the best PSF, the method employed for modeling the galaxy, and the potential biases of the fitting pipeline. Galaxy simulations are invaluable tools for understanding the performance of quantitative fitting pipelines because they provide control over the aforementioned factors (e.g., Trujillo et al. 2007; Cimatti et al. 2008; Mancini et al. 2010; Szomoru et al. 2010, 2012;

van Dokkum & Brammer 2010; Williams et al. 2010; Papovich et al. 2012; and van der Wel et al. 2012 for high- $z$  galaxies and Häussler et al. 2007 and Meert et al. 2013 for low- $z$  galaxies).

By taking advantage of existing observations, detailed analysis, and the multi-component picture of the local elliptical galaxies from H13, we address two key questions:

1) Do the quiescent massive galaxies at  $z = 2$  have sizes comparable with the quiescent massive local galaxies and therefore their observed compactness is not intrinsic but an artifact of inappropriate and insufficient morphological modeling?

2) What are the true uncertainties of the size and total luminosity measurements at different  $S/N$  levels due to structural complexities?

Two sets of simulations are performed to address these questions. Model galaxies with a single Sérsic component with parameter ranges that cover the *HST*/WFC3 observations of the red nuggets are the first set of simulations we perform. These idealized simulations can determine the robustness of fitting pipelines at different noise levels (Häussler et al. 2007). Therefore, the uncertainties and systematic errors measured from these simulated galaxies can be considered as lower limits. More informative analysis is done by simulating multi-component galaxies with the properties obtained from H13. By artificially scaling modern-day galaxies to sizes and luminosities comparable to those found in galaxies at  $z \approx 2$ , we can fit them with single-component Sérsic models to understand the systematics caused by complexities in galaxy structures. The differences between the input and output parameters shed light on possible uncertainties and biases in size and total luminosity measurements of the high- $z$  galaxies.

This paper is organized as follows. Section 2 gives the details of the galaxy simulations used throughout this study. In Section 3, we present the main results of our **GALFIT** models. In Section 4, we compare our results with similar studies in the literature. Implications for red nuggets are discussed in Section 5, ending with a summary in Section 6.

All the results assume a standard cosmology ( $H_0 = 71 \text{ km}^{-1} \text{ s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.27$ , and  $\Omega_\Lambda = 0.73$ ) and AB magnitudes.

## 2. SIMULATIONS

Simulations can determine the robustness of fitting pipelines like **GALFIT**<sup>7</sup>. In this section, we describe the different sets of simulations in detail.

Two sets of objects are simulated in order to address different issues. The first, using a single-component Sérsic profile, is to establish the baseline capability due only to  $S/N$ , under the most idealized situations, and in the absence of any complexities. This sets the fundamental limits of the technique. The second set of simulations, by rescaling models of nearby elliptical galaxies and all their sub-components to comparable sizes and  $S/N$  found at high  $z$  and fitting them using a single-component Sérsic model, directly tests the null hypothesis that early-type galaxies have not evolved in size since  $z \approx 2$ . We use Ned Wright’s cosmology calculator<sup>8</sup> to compute redshift-dependent quantities.

<sup>6</sup> <http://cgs.obs.carnegiescience.edu/CGS/Home.html>

<sup>7</sup> <http://users.obs.carnegiescience.edu/peng/work/galfit/galfit.html>

<sup>8</sup> <http://www.astro.ucla.edu/~wright/CosmoCalc.html>

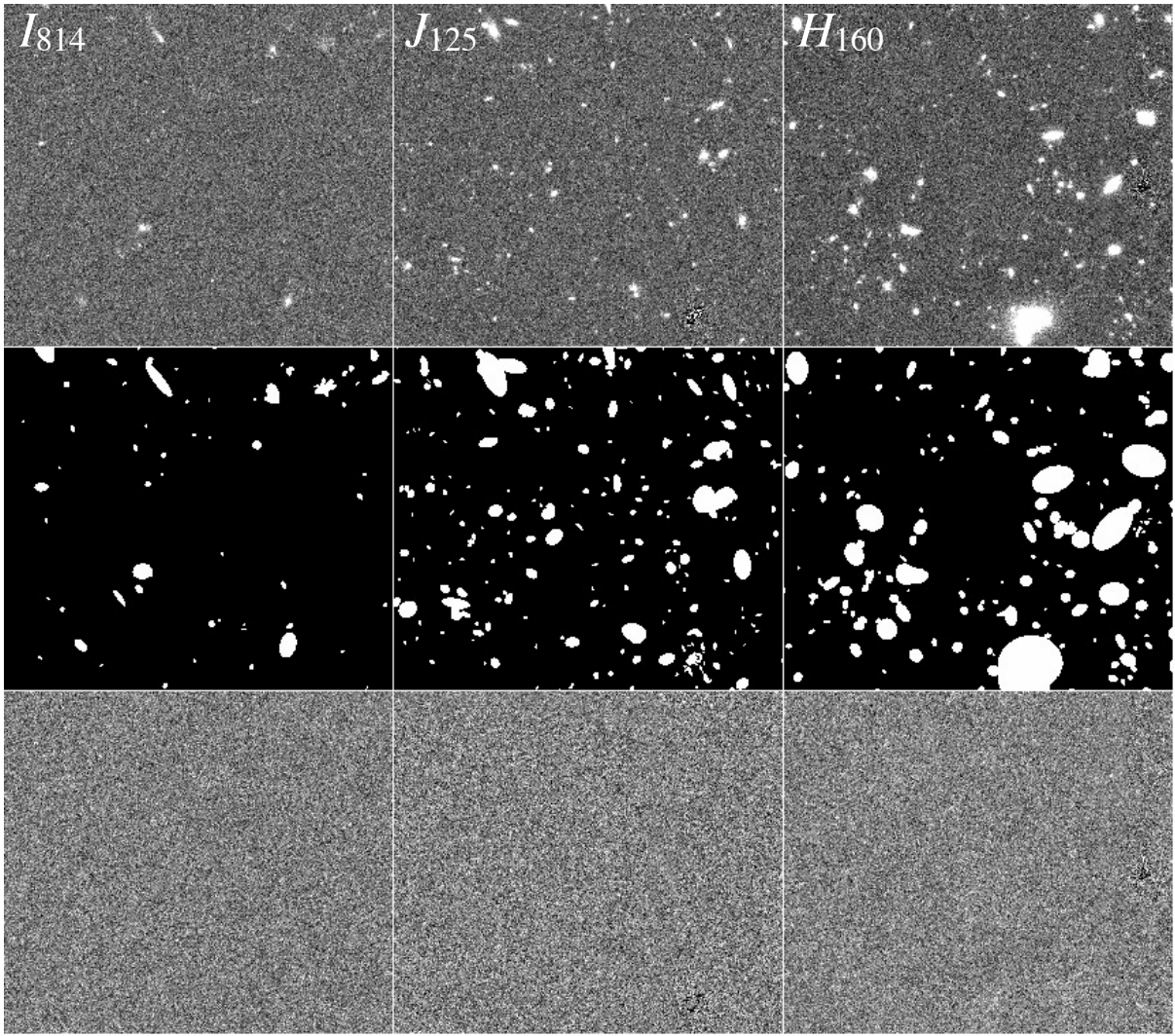


FIG. 1.— Sky backgrounds drawn from CANDELS UDS mosaic images in  $I_{814}$ ,  $J_{125}$ , and  $H_{160}$ . The simulated galaxies are put at random positions on these backgrounds. The regions are  $1021 \times 1021$  pixels (i.e.,  $\sim 30''$  in  $I_{814}$  and  $\sim 60''$  in  $J_{125}$  and  $H_{160}$ ), which is the size of the *HST*/ACS and *HST*/WFC3 CCDs. The middle panels show the bad pixel mask images of each sky background; the bottom panels are the backgrounds after removing the undesired objects. In the bottom panels, the masked pixels are replaced by randomly chosen background pixels. The bad pixel masks are used for GALFIT modeling.

### 2.1. Method

We use **GALFIT 3.0** for the simulations. **GALFIT** is an image analysis algorithm that can model profiles of galaxies, stars, and other astronomical objects in digital images. If successful, the features of interest are summarized into a small set of numbers, such as size, luminosity, profile central concentration, and geometrical parameters. **GALFIT** uses several common functions in the astronomical literature, including: exponential, Sérsic/de Vaucouleurs, Nuker, Gaussian, King, and Mofat. Out of all the functions, we exploit only the Sérsic (1968) profile and the sky background component. The sky component fits the background with a plane with a constant slope and therefore can correct for any non-flatness to first order. The Sérsic component describes the radial surface brightness profile of a galaxy as

$$\Sigma(R) = \Sigma_e \exp \left\{ -\kappa \left[ \left( \frac{R}{R_e} \right)^{1/n} - 1 \right] \right\}, \quad (1)$$

where  $R_e$  is the effective radius of the galaxy (equivalent to  $r_{50}$ , the half-light radius),  $\Sigma_e$  is the surface brightness at  $R_e$ , the Sérsic index  $n$  describes the profile shape, and the parameter  $\kappa$  is closely connected to  $n$  (Ciotti 1991). Together with position ( $x$  and  $y$ ), axis ratio  $b/a$ , and position angle, this profile has seven free parameters. The Sérsic profile represents a more general form of the exponential light profile seen in galactic disks ( $n = 1$ ; Freeman 1970) and the  $R^{1/4}$ -law (de Vaucouleurs law; de Vaucouleurs 1948) typical of luminous early-type galaxies ( $n = 4$ ). Modeling with this profile has been explored in detail in several works (e.g., Simard & Pritchet 1998; Simard et al. 2002; Graham et al. 2005). Many authors have used a constant value of  $n = 2.0$  or  $2.5$  to crudely distinguish early-type (bulge-dominated) from late-type (disk-dominated) galaxies (e.g., Blanton et al. 2003; Shen et al. 2003; Bell et al. 2004; Hogg et al. 2004; Ravindranath et al. 2004; Barden et al. 2005; McIntosh et al. 2005; Fisher & Drory 2008). Sérsic profiles with higher Sérsic indices have longer tails that make the analysis of

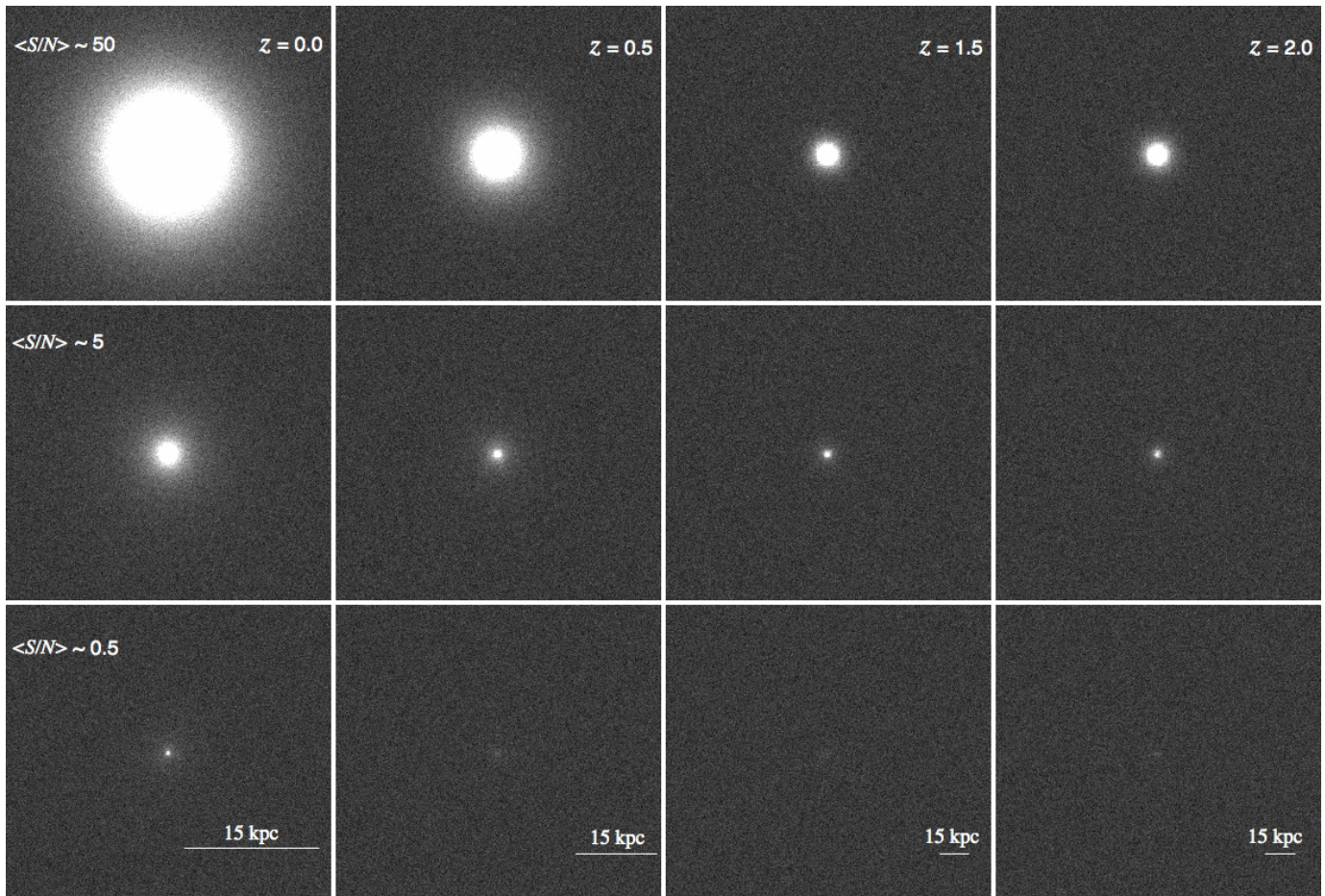


FIG. 2.— Simulated multi-component galaxies at four different redshifts and three averaged signal-to-noise ratios ( $\langle S/N \rangle$ ). The model is based on the best-fit model for NGC 1379 from H13. The simulated galaxy at  $z = 2.0$  (right columns) has  $R_e = 6$  pixels and  $e = 0$  and therefore this galaxy is shown at  $S/N \approx 5, 50$ , and  $500$ .

these galaxies more challenging due to greater sensitivity to neighboring objects, to profile mismatch, and to accurate knowledge of the sky background (see Figure 3 of Peng et al. 2010). We will compare our results for different ranges of Sérsic indices.

Sky background and noise are added to all the images in two different ways: (1) adding an artificial background level with Poisson noise using *IRAF/mknoise*<sup>9</sup> (Tody 1986, 1993); (2) putting the simulated galaxy on an actually observed background. Although the main focus of this study is galaxies at  $z = 2$ , our simulations are more generally applicable. The simulations explore a range of fundamental measurables, i.e. sizes, Sérsic indices, axis ratio, and  $S/N$ . As the simulations extend down to instrumental limits, they apply to galaxies at any redshifts provided that one first converts apparent magnitude into  $S/N$ . As illustration, we compare galaxies at  $z = 0.5, 1.5$ , and  $2$ , to our simulations. The simulated observed backgrounds are taken from CANDELS<sup>10</sup> (Grogin et al. 2011; Koekemoer et al. 2011) UDS (UKIDSS Ultra-Deep Survey; Lawrence et al. 2007)

$I_{814}$ ,  $J_{125}$ , and  $H_{160}$  mosaic images, respectively (Figure 1). Therefore, when local galaxies are rescaled to specifically match in size and  $S/N$  of observed galaxies in those filters,  $K$ -correction is not required. The simulated backgrounds resemble the observed backgrounds as they have identical gain, pixel scale, and magnitude zero point ( $mag_{zpt}$ ) as in the UDS mosaic images. Furthermore, the background root-mean-square (RMS; the standard deviation from the median sky value) of the simulated and observed backgrounds are comparable.

All artificially generated galaxies are convolved with a CANDELS UDS hybrid PSF (van der Wel et al. 2012) of the filter that corresponds to the redshift of interest. An important factor in obtaining a best-fit model is the accuracy of the PSF. The effective radii of the red nuggets are comparable to and in some cases smaller than the PSF full-width at half maximum (FWHM). Therefore, one may expect some offset in *GALFIT* measurements when an inaccurate PSF is used. This issue is studied in the Appendix, where we show that the effects of using slightly different PSFs are not more than 5% on the final model.

One of the most important factors in morphology analysis is  $S/N$ . There are several ways to define  $S/N$  because galaxies are extended. One way, which is somewhat analogous to total  $S/N$ , is given by

<sup>9</sup> IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

<sup>10</sup> <http://candels.uchicago.org/>



$$S/N = \frac{f_{\text{galaxy}}}{\sqrt{f_{\text{galaxy}} + A\sigma^2}}, \quad (2)$$

where three different parameters determine the  $S/N$ : the area  $A$  of the aperture in which the  $S/N$  is measured, the galaxy total flux  $f_{\text{galaxy}}$  within  $A$ , and the background RMS  $\sigma$ . The RMS here is the sum of all possible sources of noise, including shot noise from the sky, readout noise, and shot noise in dark current. In this paper,  $A$  includes every pixel within  $R_e$ . Note that this is different from measuring  $S/N$  within an aperture with a constant size because the aperture size here is different for each galaxy. Equation 2, however, may not be the most informative. For various technical reasons, it is harder to measure structural properties of low-surface brightness galaxies. Thus, another useful definition is analogous to an average surface brightness, defined as  $\langle S/N \rangle = \frac{S/N}{\sqrt{\pi R_e^2 (1-e)}}$ .

Here, the ellipticity  $e = 1 - b/a$ . Figure 2 shows a set of galaxies and the corresponding  $S/N$  definitions.

By adding galaxies to an observed background, one can study the effect of faint undetected objects or being in the vicinity of bright and/or extended objects. One way to minimize the effects of neighboring objects is to mask them out. We accomplish this by running SExtractor<sup>11</sup> (Bertin & Arnouts 1996) with the parameter *Detection\_threshold* set to 1.5. Instead of using the SExtractor segmentation map, the location (XPEAK\_IMAGE, YPEAK\_IMAGE), ellipticity, position angle, and the area of the objects from the SExtractor output catalog are used to create a bad pixel mask image. First, the semi-major axis of each galaxy is calculated by using the measured area and ellipticity of the object. Bad pixel regions are constructed by assigning the center, the semi-major axis, the semi-minor axis, and the position angle of each object. We double the calculated semi-major and semi-minor axes (i.e., quadruple the area) in order to mask out most of the faint outer parts of the objects. Figure 1 shows the observed backgrounds and bad pixel masks. These masks are used for GALFIT modeling. While masking does not fully eliminate the possibility of neighboring contamination (see Häussler et al. 2007), this method is analogous to that of some studies in the literature and is thus informative. Together with analysis on idealized backgrounds, these two scenarios give an idea of the different levels of systematics.

## 2.2. Single-component Galaxies

Simulating single-component galaxies is one of the most basic, yet important ways of testing the robustness of fitting pipelines (Häussler et al. 2007). We simulate more than 10,000 single-component galaxies. The Sérsic component parameters are randomly chosen from the following ranges:  $0.8 < R_e < 35$  pixels ( $\sim 0''.05$ – $2''.0$  in WFC3/ $H_{160}$ ),  $0.5 < n < 5$ , and  $0 < e < 0.8$ . For one *HST* orbit, a  $S/N$  range between 1 and 1000 within the effective radius of the simulated galaxies roughly corresponds to a magnitude range of  $20 < m_H < 27$ .

Once the galaxies are simulated, we fit each with a single Sérsic component plus a sky component. For the galaxy component, the initial parameter guesses include

the position of the galaxy (physical  $x$  and  $y$ ), Sérsic index, magnitude, effective radius, ellipticity, and position angle.

## 2.3. Multi-component Galaxies

Real galaxies have structures that complicate the analysis. Several recent studies looked into this issue by creating bulge + disk models, and then fitting single-component Sérsic models to them (Meert et al. 2013; Mosleh et al. 2013). The question was whether the disk components were missed by the analysis technique, therefore leading to underestimation of the sizes. They found that single-component models were able to recover the magnitude and sizes, with only small systematic errors (0%–20%), and that the multi-component analysis performed better.

While those studies are informative in their own right, there is another complementary approach. Our simulations are designed to more directly test the null hypothesis that elliptical galaxies today are the direct descendants of  $z \approx 2$  galaxies that have undergone no morphological changes in structure or size. To do so, we start with 100 nearby elliptical galaxies with multi-component decomposition models from H13 that capture the detailed structures down to very low surface brightnesses, accounting for the PSF. We then rescaled the composite (sum of the multiple components) as a single unit to span a range of sizes and signal-to-noise similar to galaxies at  $z \approx 2$ . We convolve the composite model with a PSF and then perform single-component Sérsic fits to it. The rationale for this approach is to be agnostic about the structural nature of high- $z$  systems (e.g., whether they are disks or bulges) because that judgment may itself be entangled with the technique as well as other more subtle technical nuances.

We rescale the sizes and luminosities of the H13 multi-component models to generate about 300 multi-component galaxies with  $R_e \approx 1 - 70$  pixels and  $S/N$  from a few to 1000. This range is relevant for studying massive early type galaxies down to the resolution and  $S/N$  limits of *HST* observations, regardless of redshift. Images of galaxies analogous to being at  $z = 0.5$ , 1.5, and 2.0 are simulated based on the  $S/N$  properties of  $I_{814}$ ,  $J_{125}$ , and  $H_{160}$  (equivalent to the rest-frame optical) CANDELS UDS images, respectively.

As in the single-component simulations, we add either a simulated background or an observed background. By using 300 multi-component galaxies as templates, more than 6000 galaxies with  $5 < S/N < 1000$  (uniformly distributed in logarithmic space) are simulated.

We analyze the multi-component galaxies in a manner similar to that used to treat the single-component galaxies. To replicate the most commonly adopted analysis method, we fit every galaxy with a single Sérsic and sky component. However, because of the mixture of different components, another method is needed to determine the intrinsic size and other parameters of the galaxy. To establish the baseline for comparison of sizes and other parameters, we employ IRAF/ellipse (Jedrzejewski 1987), a task that fits elliptical isophotes to images, to create a curve of growth (CoG) of the cumulative radial flux distribution of the model galaxy. This is generated with an ellipticity and position angle fixed to the radially average value. No noise is added, and PSF convolution is

<sup>11</sup> <http://www.astromatic.net/software/sextractor>

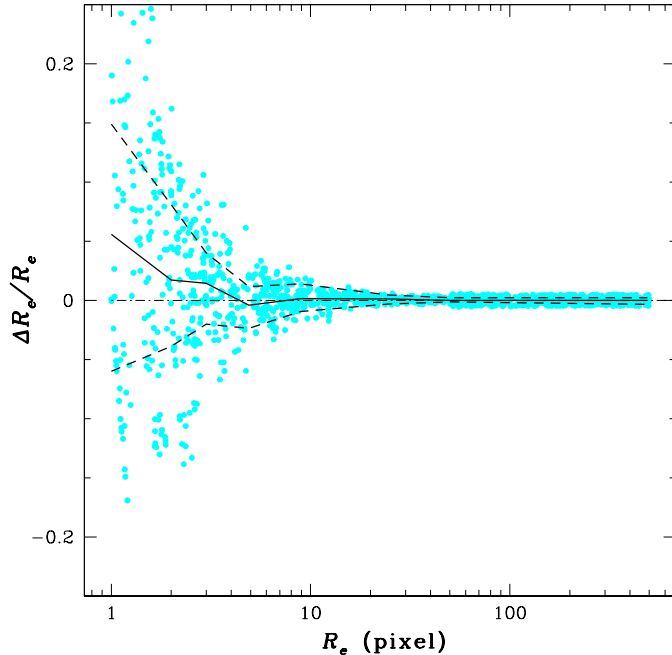


FIG. 3.— Testing the reliability of size measurements using curve of growth (CoG) analysis. The  $x$ -axis shows the effective radius of 1250 single-component galaxies, and the  $y$ -axis shows the difference between the actual size and the size measured by CoG. Black solid and dashed lines indicate the median and  $1\sigma$  uncertainties. The dot-dash line shows zero offset. The size of galaxies with  $1 < R_e < 2$  pixels can be overestimated by 5% and with 10% uncertainties. For galaxies with  $2 < R_e < 5$  pixels, the sizes can be measured without any bias and with 5% uncertainties. And for larger galaxies the uncertainties are less than 1% with no bias.

not applied. As the total flux of the galaxy is known, the CoG allows us to obtain a non-parametric estimate of the effective radius, the radius at which half of the total flux is enclosed.

To examine the robustness of our CoG analysis, we employ this method to measure the sizes of 1250 single-component galaxies with a large range of sizes (1–500 pixels). Figure 3 shows that there is no bias in size measurements for the larger galaxies but the finite size of the pixels makes size determination of the smallest galaxies slightly uncertain. Sizes with  $1 < R_e < 2$  pixels can be overestimated by 5% and with 10% uncertainty. For galaxies with  $2 < R_e < 5$  pixels, the sizes can be measured without any bias and with 5% uncertainty. And for larger galaxies the uncertainties are less than 1% with no bias. We take these effects into consideration when analyzing the multi-component galaxies. However, fewer than 2% of the simulated galaxies have  $R_e < 3$  pixels.

Lastly, we note that because the CoG technique averages over all elliptical annuli to measure the effective radius, the comparison between CoG and single-component Sérsic fits in principle is imperfect for multi-component galaxies whose isophotes often change in ellipticity with radius. Nevertheless, the general agreement in the size measurements between the two methods is reassuring.

### 3. RESULTS

#### 3.1. Single-component Galaxies

Figure 4 and Table 1 summarize the results of single-component galaxies with a simulated background; these are the most idealized simulations to establish the baseline behavior of the analysis technique for the  $S/N$  range of interest. Black solid and dashed lines in the scatter plots indicate the median and  $1\sigma$  uncertainties of different measurements. Cyan points and cyan solid lines show the results of galaxies with input  $n < 2.5$ , and red points and red dotted lines are for galaxies with  $n \geq 2.5$ .

To be consistent with the multi-component models later on, where effective radius is more ambiguous to define, the analyzed effective radii are circularized effective radii,  $R_{e,\text{circularized}} \equiv R_{e,\text{GALFIT}} \times \sqrt{1-e}$ .

We see no systematic errors in size, total luminosity, Sérsic index, or ellipticity measurements down to  $m_H = 26$ . And even at  $26 < m_H < 27$  the systematic errors are less than  $\sim 10\%$ . As expected, the uncertainties increase rapidly with decreasing  $S/N$  and particularly for  $m_H > 26$ . The Sérsic index is most vulnerable to large uncertainties. The simulated images resemble single-orbit *HST*/WFC3  $H_{160}$  observations; a galaxy with  $R_e = 5$  pixels (i.e.,  $0''.3$  or 2.5 kpc at  $z = 2$ ),  $e = 0$ , and  $m_H = 26$  has  $S/N \approx 15$  within  $R_e$ .

The uncertainties are, on average, higher for galaxies with larger Sérsic indices but the systematic offsets are comparable at different  $n$  (note that all the simulated single-component galaxies have  $n < 5$ ). Most of the galaxies with the largest scatter within a magnitude range are ones with the largest sizes, as they have lower mean surface brightness compared to the smaller galaxies.

Figure 5 and Table 2 summarize the results of single-component galaxies with an observed background, derived from real images. The scatter is larger but systematic errors are still absent down to  $m_H \approx 25$ . One of the main sources of the increased scatter is the presence of bright and/or extended objects in the actual images of the observed background (Figure 1).

It is worth pointing out that masking out objects (using a bad pixel mask for GALFIT fitting) is not the most effective way to mitigate the effects of neighboring objects. Instead, one needs to fit the target galaxy and its neighboring objects simultaneously (e.g., Häussler et al. 2007; Barden et al. 2012). The properties (e.g., magnitude, size, and Sérsic index) of the neighboring objects can affect the fit of the target object. In particular, neighbors with the largest  $n \times R_e$  have the largest effects.

Another important factor to consider is that we do not provide an input sigma image (i.e., noise map) for GALFIT; instead, we allow GALFIT to calculate it. CANDELS images are generated through an extensive process from raw, single exposures to final, drizzled mosaics. The noise tends to be correlated after the `multidrizzle` procedure, which may lead to underestimating the noise in actual data (Fruchter et al. 2009). The sigma image generated by GALFIT may not include all the necessary information about the image characteristics required for a faithful noise map. van der Wel et al. (2012) have analyzed the CANDELS UDS  $H_{160}$  image and note that the total flux of objects with  $m_H \approx 22 - 23$  and the size of typical objects in their sample (i.e.,  $\sim 0''.3$  or  $\sim 5$  pixels) correspond to the typical background flux level measured within the effective radius. Although our im-

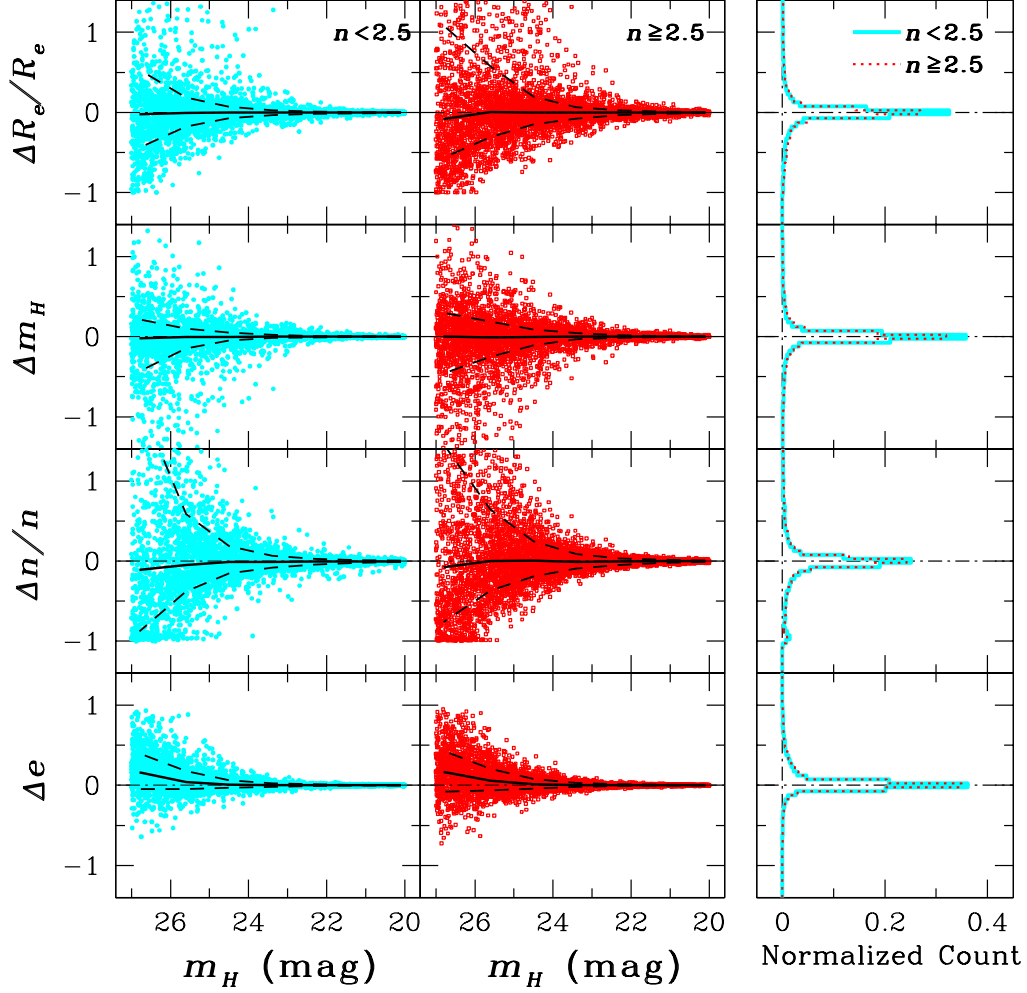


FIG. 4.— The results of more than 10,000 single-component model galaxies with a simulated background. From top to bottom, panels show the offsets between the measured and the actual effective radius  $R_e$ , magnitude  $m_H$  (in one *HST* orbit), Sérsic index  $n$ , and ellipticity  $e$ , respectively. Black solid and dashed lines in the scatter plots indicate the median and  $1\sigma$  uncertainties of different measurements. The dot-dash lines show the zero offset. Cyan points and cyan solid lines show the results of galaxies with input  $n < 2.5$ , and red points and red dotted lines are for galaxies with  $n \geq 2.5$ . The median offsets are zero down to  $m_H \approx 26$ , but the scatter rises steeply in the magnitude range between 25 and 26.

TABLE 1  
SINGLE-COMPONENT GALAXIES WITH SIMULATED BACKGROUND

$m_H$ (1)	$\Delta R_e/R_e$		$\Delta m_H$		$\Delta n/n$		$\Delta e$	
	$n < 2.5$ (2)	$n \geq 2.5$ (3)	$n < 2.5$ (4)	$n \geq 2.5$ (5)	$n < 2.5$ (6)	$n \geq 2.5$ (7)	$n < 2.5$ (8)	$n \geq 2.5$ (9)
20 – 21	$-0.00^{+0.01}_{-0.01}$	$-0.00^{+0.01}_{-0.01}$	$0.00^{+<0.01}_{-<0.01}$	$0.00^{+0.01}_{-<0.01}$	$-0.0^{+<0.1}_{-<0.1}$	$-0.0^{+<0.1}_{-<0.1}$	$0.00^{+<0.01}_{-<0.01}$	$0.00^{+<0.01}_{-<0.01}$
21 – 22	$-0.00^{+0.01}_{-0.01}$	$-0.00^{+0.01}_{-0.02}$	$0.00^{+0.01}_{-<0.01}$	$0.00^{+0.01}_{-0.01}$	$-0.0^{+<0.1}_{-<0.1}$	$-0.0^{+<0.1}_{-<0.1}$	$-0.00^{+0.01}_{-0.01}$	$0.00^{+0.01}_{-0.01}$
22 – 23	$-0.00^{+0.01}_{-0.01}$	$-0.00^{+0.03}_{-0.03}$	$0.00^{+0.01}_{-0.01}$	$0.00^{+0.02}_{-0.01}$	$-0.0^{+<0.1}_{-<0.1}$	$-0.0^{+<0.1}_{-<0.1}$	$0.00^{+0.01}_{-0.01}$	$0.00^{+0.01}_{-0.01}$
23 – 24	$-0.00^{+0.03}_{-0.03}$	$-0.00^{+0.07}_{-0.07}$	$0.00^{+0.02}_{-0.02}$	$0.00^{+0.04}_{-0.03}$	$-0.0^{+0.1}_{-0.1}$	$-0.0^{+0.1}_{-0.1}$	$0.00^{+0.02}_{-0.02}$	$0.00^{+0.03}_{-0.02}$
24 – 25	$-0.00^{+0.07}_{-0.07}$	$-0.00^{+0.20}_{-0.15}$	$-0.00^{+0.05}_{-0.05}$	$-0.00^{+0.09}_{-0.11}$	$-0.0^{+0.2}_{-0.1}$	$0.0^{+0.2}_{-0.2}$	$0.01^{+0.06}_{-0.04}$	$0.02^{+0.06}_{-0.06}$
25 – 26	$-0.01^{+0.19}_{-0.17}$	$0.00^{+0.57}_{-0.31}$	$-0.01^{+0.10}_{-0.14}$	$-0.01^{+0.19}_{-0.25}$	$-0.1^{+0.6}_{-0.3}$	$-0.0^{+0.6}_{-0.4}$	$0.04^{+0.13}_{-0.09}$	$0.06^{+0.15}_{-0.12}$
26 – 27	$-0.02^{+0.55}_{-0.42}$	$-0.08^{+1.17}_{-0.48}$	$-0.02^{+0.24}_{-0.41}$	$0.00^{+0.30}_{-0.43}$	$-0.1^{+2.1}_{-0.8}$	$-0.1^{+1.5}_{-0.7}$	$0.16^{+0.24}_{-0.21}$	$0.17^{+0.26}_{-0.25}$

NOTE. — Col. (1)  $H_{160}$  magnitude. Col. (2)-(3) The effective radius offset. Col. (4)-(5) The magnitude offset. Col. (6)-(7) The Sérsic index offset. Col. (8)-(9) The ellipticity offset. All the offsets are shown for input Sérsic indices smaller and larger than 2.5.

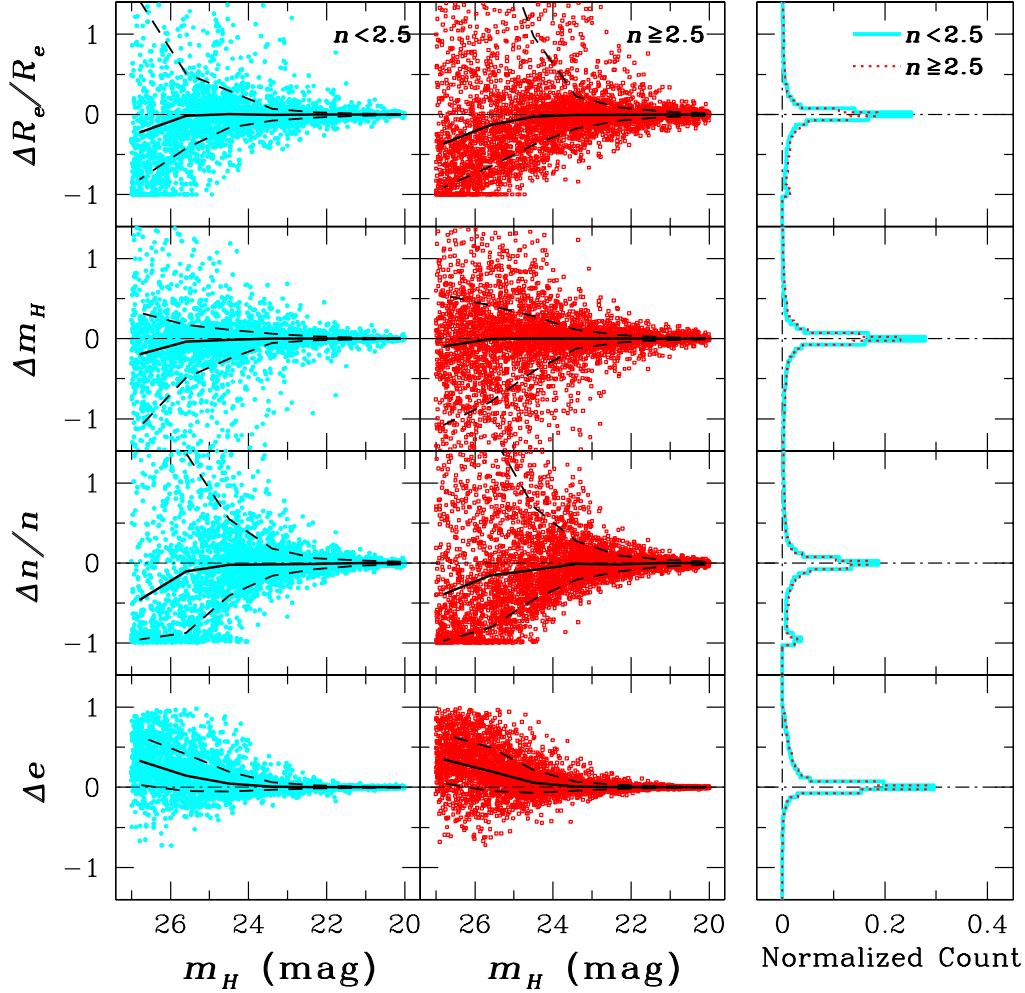


FIG. 5.— The results of more than 10,000 single-component model galaxies with an observed background. From top to bottom, panels show the offsets between the measured and the actual effective radius  $R_e$ , magnitude  $m_H$  (in one *HST* orbit), Sérsic index  $n$ , and ellipticity  $e$ , respectively. Black solid and dashed lines in the scatter plots indicate the median and  $1\sigma$  uncertainties of different measurements. The dot-dash lines show the zero offset. Cyan points and cyan solid lines show the results of galaxies with input  $n < 2.5$ , and red points and red dotted lines are for galaxies with  $n \geq 2.5$ . The median offsets are zero down to  $m_H \approx 25$ , but the scatter rises steeply in the magnitude range between 24 and 25.

TABLE 2  
SINGLE-COMPONENT GALAXIES WITH OBSERVED BACKGROUND

$m_H$ (1)	$\Delta R_e/R_e$		$\Delta m$		$\Delta n/n$		$\Delta e$	
	$n < 2.5$ (2)	$n \geq 2.5$ (3)	$n < 2.5$ (4)	$n \geq 2.5$ (5)	$n < 2.5$ (6)	$n \geq 2.5$ (7)	$n < 2.5$ (8)	$n \geq 2.5$ (9)
20 – 21	$-0.00^{+0.01}_{-0.01}$	$-0.00^{+0.01}_{-0.01}$	$0.00^{+<0.01}_{-<0.01}$	$0.00^{+0.01}_{-0.01}$	$-0.0^{+<0.1}_{-<0.1}$	$-0.0^{+<0.1}_{-<0.1}$	$0.00^{+<0.01}_{-<0.01}$	$0.00^{+<0.01}_{-<0.01}$
21 – 22	$-0.00^{+0.01}_{-0.02}$	$-0.00^{+0.04}_{-0.04}$	$-0.00^{+0.01}_{-0.01}$	$0.00^{+0.02}_{-0.02}$	$-0.0^{+<0.1}_{-<0.1}$	$-0.0^{+<0.1}_{-<0.1}$	$0.00^{+0.01}_{-0.01}$	$0.00^{+0.01}_{-0.01}$
22 – 23	$-0.00^{+0.03}_{-0.04}$	$-0.01^{+0.10}_{-0.07}$	$0.00^{+0.03}_{-0.02}$	$0.00^{+0.04}_{-0.05}$	$-0.0^{+0.1}_{-0.1}$	$-0.0^{+0.1}_{-0.1}$	$0.00^{+0.02}_{-0.02}$	$0.00^{+0.03}_{-0.02}$
23 – 24	$-0.01^{+0.08}_{-0.07}$	$-0.01^{+0.23}_{-0.17}$	$0.00^{+0.06}_{-0.05}$	$0.00^{+0.11}_{-0.13}$	$-0.0^{+0.2}_{-0.1}$	$-0.0^{+0.3}_{-0.2}$	$0.01^{+0.05}_{-0.04}$	$0.01^{+0.06}_{-0.05}$
24 – 25	$0.01^{+0.29}_{-0.17}$	$-0.03^{+1.00}_{-0.35}$	$-0.01^{+0.12}_{-0.24}$	$0.00^{+0.28}_{-0.39}$	$-0.0^{+0.6}_{-0.4}$	$-0.1^{+0.8}_{-0.4}$	$0.04^{+0.15}_{-0.10}$	$0.06^{+0.16}_{-0.13}$
25 – 26	$-0.02^{+0.53}_{-0.41}$	$-0.13^{+2.54}_{-0.52}$	$-0.04^{+0.21}_{-0.45}$	$-0.00^{+0.41}_{-0.75}$	$-0.1^{+1.5}_{-0.8}$	$-0.2^{+1.8}_{-0.7}$	$0.15^{+0.27}_{-0.19}$	$0.20^{+0.30}_{-0.24}$
26 – 27	$-0.22^{+1.65}_{-0.59}$	$-0.36^{+1.93}_{-0.55}$	$-0.20^{+0.52}_{-0.91}$	$-0.10^{+0.64}_{-0.88}$	$-0.5^{+3.9}_{-0.5}$	$-0.4^{+2.7}_{-0.6}$	$0.33^{+0.30}_{-0.30}$	$0.35^{+0.31}_{-0.29}$

NOTE. — Col. (1)  $H_{160}$  magnitude. Col. (2)-(3) The effective radius offset. Col. (4)-(5) The magnitude offset. Col. (6)-(7) The Sérsic index offset. Col. (8)-(9) The ellipticity offset. All the offsets are shown for input Sérsic indices smaller and larger than 2.5.



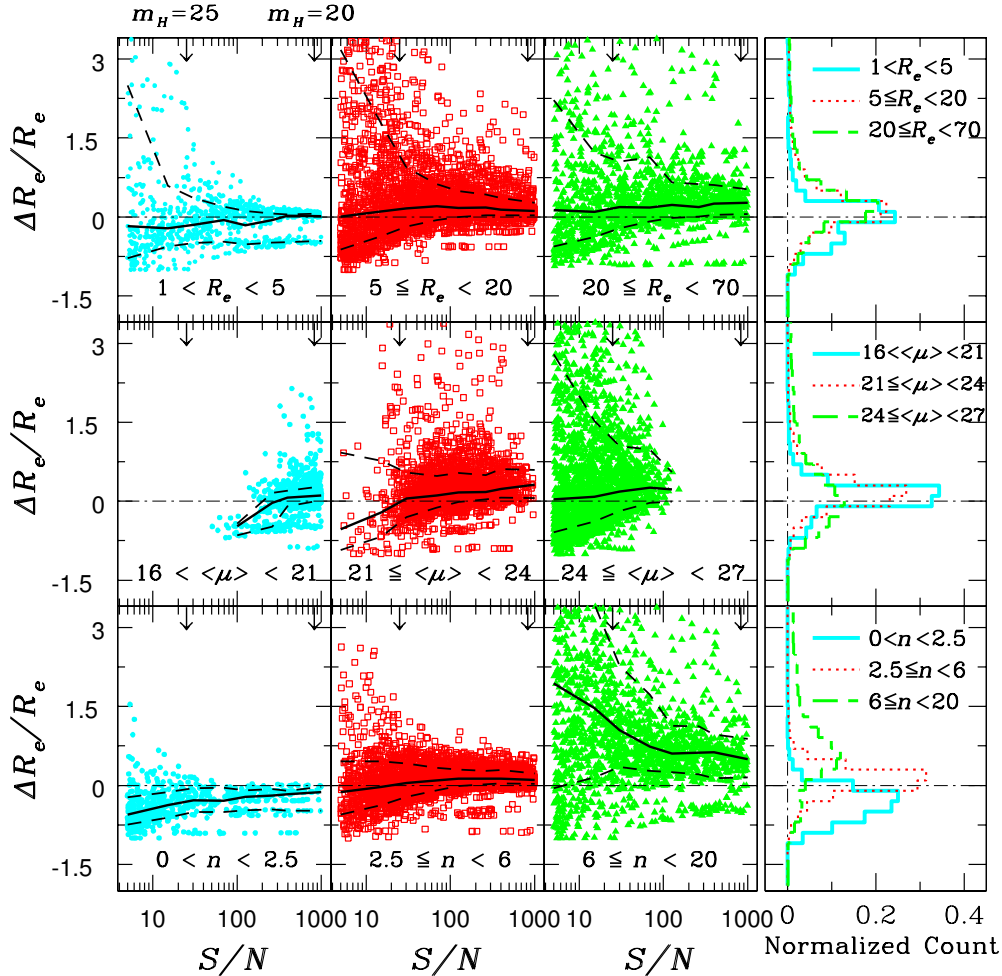


FIG. 6.— Results of more than 7,000 multi-component galaxies with simulated backgrounds. The upper, middle, and lower panels show the results at different intervals of size  $R_e$  (pixel), mean surface brightness  $\langle \mu \rangle$  ( $m_H$  arcsec $^{-2}$ ), and Sérsic index  $n$ , respectively. Black solid and dashed lines in the scatter plots indicate the median and  $1\sigma$  uncertainties of different measurements. The dot-dash lines show the zero offset. The downward-pointing arrows in the scatter plots indicate the  $S/N$  of a galaxy with  $R_e = 5$  pixels,  $e = 0$ , and  $m_H = 20$  and 25. Note the significant size and total luminosity overestimations of GALFIT models with  $n > 6$ ; see text in Section 3.2 regarding the offsets in different bin internals of Sérsic index.

ages with simulated backgrounds have similar RMS to CANDELS UDS  $H_{160}$  image, the galaxy flux and the background flux are comparable at  $m_H \approx 23.5$  for images with simulated backgrounds. This can explain some of the discrepancies between Figure 4 and Figure 5, even at the bright end.

### 3.2. Multi-component Galaxies

Simulated galaxies with multiple components, where the subcomponents are taken from accurate decompositions of nearby galaxies of H13, provide a better description of the observed galaxy structures than idealized single-component models. We examine the reliability of measuring the size and total luminosity of multi-component galaxies by single-component fitting.

Figure 6 and Tables 3 and 4 show the difference between actual and measured effective radii at different  $S/N$  ranges for about 7,000 model galaxies in images with simulated (i.e. idealized) backgrounds. The upper, middle, and lower panels give the results at different intervals of size, mean surface brightness, and Sérsic index,

respectively. Black solid and dashed lines in the scatter plots indicate the median and  $1\sigma$  uncertainties of different measurements. The downward-pointing arrows on the top of each subpanel indicate the  $S/N$  of a galaxy with  $R_e = 5$  pixels,  $e = 0$ , and  $m_H = 20$  or 25 (one orbit). Typical red nugget galaxies in CANDELS have  $m_H \approx 20 - 23$ .

A notable feature in Figure 6 are the results for galaxies with  $6 < n < 20$ . Note that we do not impose an upper limit on the Sérsic index;  $n = 20$  is an internal upper limit set by GALFIT. It is clear that the fits with  $n > 6$  are generally unreliable. Sérsic profiles with  $n > 4$  have long tails. In a regime where the surface brightness of the galaxy is close to the sky level, the flux in the long tail can be overestimated due to the inherent degeneracy between a Sérsic profile with large  $n$  and any residual background flux. This leads not only to overestimating the total flux of the galaxy but also its size. On the other hand, for all practical purposes, one can correct for this effect by refitting the galaxies with  $n > 6$  by fixing  $n$  to 6. Figure 7 demonstrates the significant improvement

TABLE 3  
MULTI-COMPONENT GALAXIES WITH SIMULATED BACKGROUND  
 $\Delta R_e/R_e$

(1)	Subsample (2)	$S/N$							
		5 – 10 (3)	10 – 20 (4)	20 – 35 (5)	35 – 70 (6)	70 – 135 (7)	135 – 265 (8)	265 – 515 (9)	515 – 1000 (10)
$R_e$	1 – 5	$-0.24^{+2.39}_{-0.49}$	$-0.21^{+1.68}_{-0.31}$	$-0.06^{+0.64}_{-0.34}$	$-0.01^{+0.31}_{-0.32}$	$0.03^{+0.19}_{-0.49}$	$0.04^{+0.09}_{-0.49}$	$0.04^{+0.06}_{-0.49}$	$0.03^{+0.06}_{-0.13}$
	5 – 20	$0.02^{+3.24}_{-0.62}$	$0.14^{+1.45}_{-0.50}$	$0.17^{+0.79}_{-0.35}$	$0.21^{+0.42}_{-0.27}$	$0.18^{+0.37}_{-0.18}$	$0.19^{+0.26}_{-0.15}$	$0.15^{+0.22}_{-0.12}$	$0.12^{+0.16}_{-0.08}$
	20 – 70	$0.13^{+2.09}_{-0.69}$	$0.09^{+1.13}_{-0.48}$	$0.19^{+0.86}_{-0.41}$	$0.18^{+0.95}_{-0.26}$	$0.23^{+0.43}_{-0.24}$	$0.19^{+0.43}_{-0.18}$	$0.25^{+0.36}_{-0.21}$	$0.27^{+0.26}_{-0.22}$
$\langle\mu\rangle$	16 – 21	...	...	...	...	$-0.48^{+0.05}_{-0.17}$	$-0.03^{+0.19}_{-0.45}$	$0.07^{+0.14}_{-0.16}$	$0.11^{+0.17}_{-0.09}$
	21 – 24	$-0.53^{+1.45}_{-0.40}$	$-0.21^{+0.98}_{-0.46}$	$0.05^{+0.49}_{-0.35}$	$0.11^{+0.37}_{-0.23}$	$0.17^{+0.37}_{-0.18}$	$0.18^{+0.32}_{-0.15}$	$0.23^{+0.38}_{-0.16}$	$0.32^{+0.28}_{-0.26}$
	24 – 27	$0.04^{+2.75}_{-0.63}$	$0.09^{+1.44}_{-0.47}$	$0.19^{+0.86}_{-0.37}$	$0.26^{+0.72}_{-0.29}$	$0.23^{+0.33}_{-0.23}$	...	...	...
$n$	0 – 2.5	$-0.55^{+0.32}_{-0.19}$	$-0.37^{+0.25}_{-0.24}$	$-0.27^{+0.22}_{-0.23}$	$-0.29^{+0.24}_{-0.21}$	$-0.21^{+0.12}_{-0.26}$	$-0.19^{+0.12}_{-0.26}$	$-0.17^{+0.06}_{-0.32}$	$-0.12^{+0.10}_{-0.36}$
	2.5 – 6	$-0.12^{+0.58}_{-0.43}$	$-0.03^{+0.49}_{-0.35}$	$0.05^{+0.37}_{-0.26}$	$0.09^{+0.24}_{-0.19}$	$0.13^{+0.17}_{-0.14}$	$0.13^{+0.15}_{-0.10}$	$0.13^{+0.17}_{-0.09}$	$0.11^{+0.12}_{-0.08}$
	6 – 20	$1.93^{+4.44}_{-1.99}$	$1.47^{+1.98}_{-1.32}$	$1.05^{+1.23}_{-0.69}$	$0.74^{+0.98}_{-0.47}$	$0.60^{+0.53}_{-0.36}$	$0.62^{+0.51}_{-0.40}$	$0.63^{+0.34}_{-0.49}$	$0.50^{+0.38}_{-0.35}$

NOTE. — Col. (1) Parameter. Col. (2) The subsamples selected based on different ranges of effective radius ( $R_e$ ), surface brightness ( $\langle\mu\rangle$ ), and Sérsic index ( $n$ ). Col. (3)-(10) Equally (in logarithmic space) spaced intervals of  $S/N$ . See text in Section 3.2 regarding the offsets in different Sérsic index.

TABLE 4  
MULTI-COMPONENT GALAXIES WITH SIMULATED BACKGROUND  
 $\Delta m_H$

(1)	Subsample (2)	$S/N$							
		5 – 10 (3)	10 – 20 (4)	20 – 35 (5)	35 – 70 (6)	70 – 135 (7)	135 – 265 (8)	265 – 515 (9)	515 – 1000 (10)
$R_e$	1 – 5	$-0.04^{+0.41}_{-0.65}$	$0.08^{+0.22}_{-0.64}$	$0.02^{+0.30}_{-0.27}$	$-0.01^{+0.20}_{-0.12}$	$-0.05^{+0.31}_{-0.07}$	$-0.03^{+0.22}_{-0.06}$	$-0.01^{+0.18}_{-0.05}$	$-0.02^{+0.07}_{-0.04}$
	5 – 20	$-0.14^{+0.56}_{-0.59}$	$-0.13^{+0.37}_{-0.37}$	$-0.10^{+0.21}_{-0.25}$	$-0.11^{+0.13}_{-0.16}$	$-0.10^{+0.09}_{-0.13}$	$-0.10^{+0.07}_{-0.12}$	$-0.09^{+0.07}_{-0.10}$	$-0.08^{+0.05}_{-0.08}$
	20 – 70	$-0.11^{+0.57}_{-0.62}$	$-0.06^{+0.33}_{-0.35}$	$-0.10^{+0.23}_{-0.25}$	$-0.09^{+0.12}_{-0.29}$	$-0.12^{+0.13}_{-0.14}$	$-0.10^{+0.09}_{-0.15}$	$-0.13^{+0.10}_{-0.12}$	$-0.14^{+0.11}_{-0.09}$
$\langle\mu\rangle$	16 – 21	...	...	...	...	$0.26^{+0.16}_{-0.26}$	$0.02^{+0.37}_{-0.11}$	$-0.05^{+0.09}_{-0.06}$	$-0.07^{+0.06}_{-0.09}$
	21 – 24	$0.14^{+0.45}_{-0.57}$	$0.08^{+0.36}_{-0.48}$	$-0.03^{+0.25}_{-0.23}$	$-0.07^{+0.14}_{-0.15}$	$-0.10^{+0.10}_{-0.14}$	$-0.10^{+0.07}_{-0.13}$	$-0.13^{+0.08}_{-0.13}$	$-0.16^{+0.12}_{-0.09}$
	24 – 27	$-0.14^{+0.55}_{-0.60}$	$-0.09^{+0.35}_{-0.40}$	$-0.12^{+0.24}_{-0.24}$	$-0.14^{+0.16}_{-0.20}$	$-0.12^{+0.10}_{-0.13}$	...	...	...
$n$	0 – 2.5	$0.40^{+0.35}_{-0.39}$	$0.29^{+0.27}_{-0.25}$	$0.27^{+0.25}_{-0.23}$	$0.27^{+0.33}_{-0.21}$	$0.25^{+0.25}_{-0.18}$	$0.22^{+0.32}_{-0.20}$	$0.20^{+0.37}_{-0.19}$	$0.15^{+0.4}_{-0.14}$
	2.5 – 6	$-0.04^{+0.41}_{-0.35}$	$-0.01^{+0.25}_{-0.24}$	$-0.03^{+0.17}_{-0.18}$	$-0.06^{+0.11}_{-0.11}$	$-0.08^{+0.08}_{-0.08}$	$-0.08^{+0.07}_{-0.06}$	$-0.08^{+0.06}_{-0.07}$	$-0.07^{+0.05}_{-0.06}$
	6 – 20	$-0.42^{+0.41}_{-0.67}$	$-0.45^{+0.34}_{-0.29}$	$-0.35^{+0.19}_{-0.20}$	$-0.28^{+0.15}_{-0.20}$	$-0.25^{+0.12}_{-0.11}$	$-0.25^{+0.14}_{-0.10}$	$-0.25^{+0.17}_{-0.08}$	$-0.21^{+0.14}_{-0.10}$

NOTE. — Col. (1) Parameter. Col. (2) The subsamples selected based on different ranges of effective radius ( $R_e$ ), surface brightness ( $\langle\mu\rangle$ ), and Sérsic index ( $n$ ). Col. (3)-(10) Equally (in logarithmic space) spaced intervals of  $S/N$ . See text in Section 3.2 regarding the offsets in different Sérsic index.

in the size determination. Comparing Figures 6 and 7 clarifies that most of the outliers in Figure 6 are galaxies with  $n > 6$ . It is worth noting that because the simulations are based on a finite sample of 100 nearby galaxy “templates,” rescaled in luminosity and size, the behavior of the scatter is only pseudo-random, which leads to horizontal striations in the scatter plots of Figure 6.

Figures 6 and 7 (and 8) show that when the objects are binned by measured Sérsic index, the scatter does not appear symmetric in the lowest bin in  $n$ . One might interpret the result as the fit missing the outer parts of the galaxies. However, that offset is deceptive because the

objects are plotted according to their measured Sérsic index rather than intrinsic index; one does not know the Sérsic index for real, multi-component galaxies until after the fit. Because measured Sérsic indices are positively correlated in a fit with measured sizes (Yoon et al. 2011), objects selected to have low measured Sérsic index typically would have smaller measured sizes. This is seen in the extreme, comparing Figures 6 and 7, for the bin where  $n > 6$ . In Figure 7, the measured sizes become much smaller when  $n > 6$  objects are constrained to  $n = 6$ . The apparent systematic bias is therefore an artifact of the Sérsic index not being an independent vari-

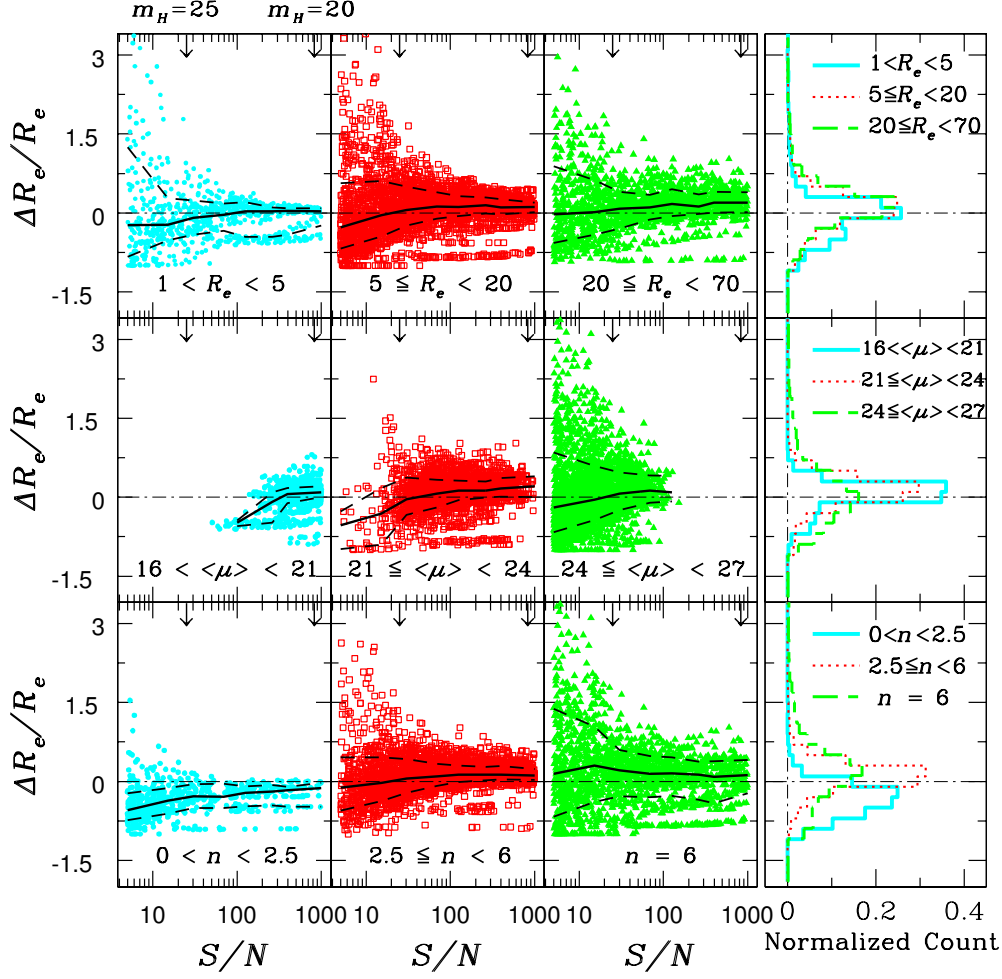


FIG. 7.— Results of more than 7,000 multi-component galaxies with a simulated background. The upper, middle, and lower panels show the results at different intervals of size  $R_e$  (pixel), mean surface brightness  $\langle \mu \rangle$  ( $m_H \text{ arcsec}^{-2}$ ), and Sérsic index  $n$ , respectively. Black solid and dashed lines in the scatter plots indicate the median and  $1\sigma$  uncertainties of different measurements. The dot-dash lines show the zero offset. The downward-pointing arrows in the scatter plots indicate the  $S/N$  of a galaxy with  $R_e = 5$  pixels,  $e = 0$ , and  $m_H = 20$  and  $25$ . Galaxies with  $n > 6$  are refit by fixing the Sérsic index to  $n = 6$ . Note the significant decrease in scatter and systematic errors (compare with Figure 6); see text in Section 3.2 regarding the offsets in different bin internals of Sérsic index.

able in the simulation. In contrast, when all the objects are grouped together without regard to Sérsic index, the scatter is more symmetric about the mean (top row).

Tables 5 and 6 summarize the results of the best-fit multi-component models. On average, across the entire  $S/N$  range of interest, the size of the galaxies and their total luminosities are slightly underestimated at very low  $S/N$ . On the other hand, a small but positive trend is present, as seen in Figure 7, at higher  $S/N$  values. The multi-component galaxies are drawn from CGS elliptical galaxies and the Sérsic indices of these galaxies should peak around  $n = 4$ , as seen in our results ( $< 15\%$  have  $n < 2.5$ , and  $< 20\%$  have  $n > 6$ ).

There is a weak size dependence in the size determination of the galaxies. The systematic offsets of galaxies with  $S/N > 50$  are smaller for smaller galaxies. Figure 3 indicates that the CoG technique to measure sizes can produce systematics even in idealized galaxies when  $R_e < 2$  pixels. Even after correcting for this effect, the systematic offsets for galaxies with  $1 < R_e < 5$  pixels

are still slightly less than what we find for galaxies with larger sizes. The slight trend and scatter compared to idealized simulations therefore illustrate the fundamental differences due to structural complexities.

Figure 8 and Tables 7 and 8 summarize the results of multi-component models with an observed sky background taken from CANDELS UDS  $H_{160}$  images. The trends are similar between images with simulated and observed backgrounds. As expected, the uncertainties are higher for images with the observed background, especially at lower  $S/N$ .

Under-subtraction or over-subtraction of the sky value can introduce spurious curvature into the brightness profile, especially in the faint, outer regions of the galaxy (e.g., MacArthur et al. 2003; Erwin et al. 2008; Bernardi et al. 2010; Yoon et al. 2011). Figure 9 shows the accuracy of GALFIT sky determination. It shows that in our simulations GALFIT measures the sky value with accuracy better than few 0.01%. Therefore, the systematic errors of single-component fits of multi-component galaxies are

TABLE 5  
MULTI-COMPONENT GALAXIES WITH SIMULATED BACKGROUND – CORRECTED FOR MODELS WITH  $n > 6$   
 $\Delta R_e / R_e$

(1)	Subsample (2)	$S/N$							
		5 – 10 (3)	10 – 20 (4)	20 – 35 (5)	35 – 70 (6)	70 – 135 (7)	135 – 265 (8)	265 – 515 (9)	515 – 1000 (10)
$R_e$	1 – 5	$-0.23^{+1.48}_{-0.60}$	$-0.22^{+0.50}_{-0.30}$	$-0.09^{+0.32}_{-0.32}$	$-0.03^{+0.21}_{-0.30}$	$0.03^{+0.17}_{-0.49}$	$0.03^{+0.08}_{-0.49}$	$0.04^{+0.05}_{-0.45}$	$0.03^{+0.06}_{-0.27}$
	5 – 20	$-0.28^{+0.85}_{-0.40}$	$-0.04^{+0.65}_{-0.40}$	$0.05^{+0.45}_{-0.30}$	$0.12^{+0.27}_{-0.24}$	$0.12^{+0.20}_{-0.17}$	$0.14^{+0.15}_{-0.15}$	$0.11^{+0.15}_{-0.14}$	$0.11^{+0.09}_{-0.10}$
	20 – 70	$-0.03^{+0.91}_{-0.54}$	$0.02^{+0.6}_{-0.43}$	$0.08^{+0.31}_{-0.38}$	$0.11^{+0.23}_{-0.23}$	$0.17^{+0.28}_{-0.23}$	$0.12^{+0.23}_{-0.14}$	$0.20^{+0.21}_{-0.19}$	$0.20^{+0.19}_{-0.19}$
$\langle \mu \rangle$	16 – 21	...	...	...	...	$-0.48^{+0.04}_{-0.07}$	$-0.09^{+0.18}_{-0.09}$	$0.06^{+0.12}_{-0.17}$	$0.09^{+0.11}_{-0.09}$
	21 – 24	$-0.53^{+0.26}_{-0.45}$	$-0.31^{+0.49}_{-0.59}$	$-0.04^{+0.41}_{-0.30}$	$0.07^{+0.26}_{-0.26}$	$0.12^{+0.19}_{-0.18}$	$0.13^{+0.18}_{-0.14}$	$0.17^{+0.20}_{-0.16}$	$0.21^{+0.19}_{-0.21}$
	24 – 27	$-0.20^{+1.05}_{-0.46}$	$-0.04^{+0.64}_{-0.39}$	$0.07^{+0.40}_{-0.31}$	$0.13^{+0.28}_{-0.23}$	$0.09^{+0.32}_{-0.17}$	...	...	...
$n$	0 – 2.5	$-0.55^{+0.32}_{-0.19}$	$-0.37^{+0.25}_{-0.24}$	$-0.27^{+0.22}_{-0.23}$	$-0.29^{+0.24}_{-0.21}$	$-0.21^{+0.12}_{-0.26}$	$-0.19^{+0.12}_{-0.26}$	$-0.17^{+0.06}_{-0.32}$	$-0.12^{+0.10}_{-0.36}$
	2.5 – 6	$-0.12^{+0.58}_{-0.43}$	$-0.03^{+0.49}_{-0.35}$	$0.05^{+0.37}_{-0.35}$	$0.09^{+0.24}_{-0.19}$	$0.13^{+0.17}_{-0.14}$	$0.13^{+0.15}_{-0.10}$	$0.13^{+0.17}_{-0.09}$	$0.11^{+0.12}_{-0.08}$
	6	$0.14^{+1.24}_{-0.81}$	$0.30^{+0.74}_{-0.70}$	$0.21^{+0.38}_{-0.49}$	$0.15^{+0.33}_{-0.45}$	$0.16^{+0.30}_{-0.43}$	$0.13^{+0.25}_{-0.48}$	$0.09^{+0.29}_{-0.50}$	$0.12^{+0.29}_{-0.34}$

NOTE. — Col. (1) Parameter. Col. (2) The subsamples selected based on different ranges of effective radius ( $R_e$ ), surface brightness ( $\langle \mu \rangle$ ), and Sérsic index ( $n$ ). Col. (3)-(10) Equally (in logarithmic space) spaced intervals of  $S/N$ . See text in Section 3.2 regarding the offsets in different Sérsic index.

TABLE 6  
MULTI-COMPONENT GALAXIES WITH SIMULATED BACKGROUND – CORRECTED FOR MODELS WITH  $n > 6$   
 $\Delta m_H$

(1)	Subsample (2)	$S/N$							
		5 – 10 (3)	10 – 20 (4)	20 – 35 (5)	35 – 70 (6)	70 – 135 (7)	135 – 265 (8)	265 – 515 (9)	515 – 1000 (10)
$R_e$	1 – 5	$-0.00^{+0.54}_{-0.69}$	$0.10^{+0.32}_{-0.30}$	$0.06^{+0.30}_{-0.20}$	$0.00^{+0.26}_{-0.12}$	$-0.04^{+0.41}_{-0.07}$	$-0.01^{+0.26}_{-0.06}$	$-0.01^{+0.33}_{-0.05}$	$-0.01^{+0.30}_{-0.05}$
	5 – 20	$0.07^{+0.47}_{-0.46}$	$-0.04^{+0.36}_{-0.27}$	$-0.05^{+0.20}_{-0.19}$	$-0.07^{+0.14}_{-0.12}$	$-0.07^{+0.09}_{-0.10}$	$-0.08^{+0.07}_{-0.07}$	$-0.07^{+0.08}_{-0.07}$	$-0.07^{+0.06}_{-0.05}$
	20 – 70	$0.03^{+0.48}_{-0.48}$	$-0.01^{+0.33}_{-0.28}$	$-0.05^{+0.26}_{-0.15}$	$-0.04^{+0.12}_{-0.12}$	$-0.09^{+0.12}_{-0.11}$	$-0.07^{+0.08}_{-0.10}$	$-0.10^{+0.09}_{-0.09}$	$-0.11^{+0.09}_{-0.07}$
$\langle \mu \rangle$	16 – 21	...	...	...	...	$0.38^{+0.07}_{-0.13}$	$0.06^{+0.33}_{-0.13}$	$-0.05^{+0.12}_{-0.05}$	$-0.06^{+0.07}_{-0.06}$
	21 – 24	$0.27^{+0.43}_{-0.28}$	$0.24^{+0.30}_{-0.40}$	$0.02^{+0.27}_{-0.20}$	$-0.05^{+0.16}_{-0.11}$	$-0.07^{+0.10}_{-0.09}$	$-0.08^{+0.07}_{-0.07}$	$-0.09^{+0.08}_{-0.09}$	$-0.11^{+0.10}_{-0.08}$
	24 – 27	$0.04^{+0.48}_{-0.49}$	$-0.02^{+0.33}_{-0.27}$	$-0.05^{+0.22}_{-0.18}$	$-0.07^{+0.13}_{-0.13}$	$-0.07^{+0.10}_{-0.13}$	...	...	...
$n$	0 – 2.5	$0.40^{+0.35}_{-0.40}$	$0.28^{+0.29}_{-0.25}$	$0.28^{+0.24}_{-0.23}$	$0.28^{+0.32}_{-0.22}$	$0.24^{+0.26}_{-0.17}$	$0.22^{+0.32}_{-0.20}$	$0.20^{+0.37}_{-0.19}$	$0.18^{+0.38}_{-0.17}$
	2.5 – 6	$-0.04^{+0.41}_{-0.34}$	$-0.01^{+0.25}_{-0.24}$	$-0.04^{+0.17}_{-0.18}$	$-0.06^{+0.11}_{-0.11}$	$-0.08^{+0.08}_{-0.08}$	$-0.08^{+0.06}_{-0.07}$	$-0.08^{+0.06}_{-0.07}$	$-0.07^{+0.05}_{-0.06}$
	6	$-0.07^{+0.48}_{-0.57}$	$-0.15^{+0.38}_{-0.28}$	$-0.10^{+0.25}_{-0.16}$	$-0.07^{+0.22}_{-0.13}$	$-0.09^{+0.21}_{-0.12}$	$-0.08^{+0.23}_{-0.09}$	$-0.06^{+0.39}_{-0.12}$	$-0.09^{+0.29}_{-0.08}$

NOTE. — Col. (1) Parameter. Col. (2) The subsamples selected based on different ranges of effective radius ( $R_e$ ), surface brightness ( $\langle \mu \rangle$ ), and Sérsic index ( $n$ ). Col. (3)-(10) Equally (in logarithmic space) spaced intervals of  $S/N$ . See text in Section 3.2 regarding the offsets in different Sérsic index.

not caused by overestimating or underestimating the sky value, but rather merely reflects the limitations of single-component modeling. In fact, even when we fix the sky value to the actual value during the fits, the outcomes do not change.

The provided tables can be used by future observers to quantify potential systematic offsets and uncertainties in their measurements of structural parameters in high- $z$  galaxies. Although the main purpose of the simulations is to determine the robustness of the GALFIT modeling of high- $z$ , compact galaxies, our simulations would be appropriate for galaxies of all redshifts, as long as

they fall within our  $S/N$ , size, Sérsic index, and parameter space. Note, however, that the multi-component simulated galaxies mimic the morphology of only early-type galaxies. A forthcoming work (R. Davari et al., in preparation) will address a broader range of morphological types.

#### 4. COMPARISON TO OTHER STUDIES

Van der Wel et al. (2012) performed comprehensive single-component simulations based on CANDELS imaging in  $H_{160}$ , based on their catalog of 6492 objects in GOODS-South. They find that  $m_H$ ,  $R_e$ , and  $e$  can be



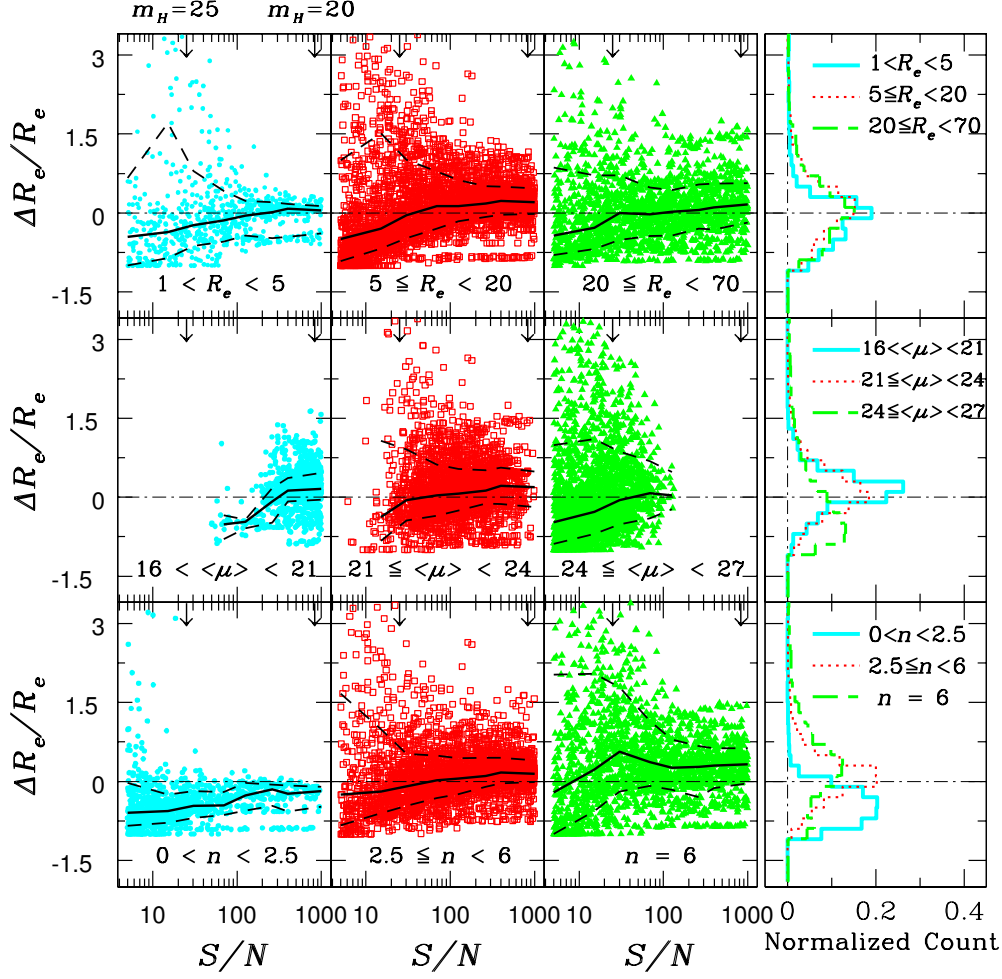


FIG. 8.— Results of more than 7,000 multi-component galaxies with an observed background. The upper, middle, and lower panels show the results at different intervals of size  $R_e$  (pixel), mean surface brightness  $\langle\mu\rangle$  ( $m_H$  arcsec $^{-2}$ ), and Sérsic index  $n$ , respectively. Black solid and dashed lines in the scatter plots indicate the median and  $1\sigma$  uncertainties of different measurements. The dot-dash lines show zero offset. The downward-pointing arrows in the scatter plots indicate the  $S/N$  of a galaxy with  $R_e = 5$  pixels,  $e = 0$ , and  $m_H = 20$  and  $25$ . Galaxies with  $n > 6$  are refit by fixing the Sérsic index to  $n = 6$ . The trends are similar to the results for simulated backgrounds, although the scatter is larger; see text in Section 3.2 regarding the offsets in different bin intervals of Sérsic index.

inferred with a random accuracy of 20% or better for galaxies brighter than  $m_H = 24.5$ , whereas  $n$  can be measured at the same level of accuracy for galaxies brighter than  $m_H = 23.5$  (their Table 3). They conclude that since a typical faint high- $z$  galaxy is small and has a low Sérsic index, 10%-level accuracy in the single-component measurements can be reached down to  $m_H \approx 24.5$ . For their faintest sources ( $m_H > 25.5$ ) with large sizes ( $R_e > 0''.4$ ), the uncertainties in the magnitude and structural parameters start to become substantial because they are dominated by the uncertainty in the background estimate. This is in agreement with our single-component results (Tables 1 and 2). We find that at a specific magnitude the scatter is higher for larger galaxies because their surface brightness is lower.

Trujillo et al. (2007) simulate 1000 single-component galaxies with properties matching the observed distribution of their objects in  $I_{814}$ . A background sky, randomly taken from the  $I_{814}$  image, is added to the generated galaxies, and the galaxy models are convolved with the

observed PSF. Based on their Figure 2, there is no systematic offset for galaxies with  $m_I < 24.0$ . For fainter galaxies, only those with  $n > 2.5$ , there is less than a 25% systematic offset toward smaller sizes. They also explore the variation of the PSF within the image to see how it can affect the recovery of the sizes. Using different stars in the image as the PSF, they find that the size estimations are robust to changes in the selected PSF; the scatter is about 10%. Abramson et al. (2013) also use six different PSFs (including an empirical PSF) and find 5%-level accuracy for lower Sérsic indices and 30%-level accuracy for higher Sérsic indices. Our Figure A2 confirms that the effects are small when using slightly inaccurate PSFs.

However, Mancini et al. (2010) claim that for objects with large effective radii and Sérsic indices, as elliptical galaxies with masses of  $2.5 \times 10^{11} M_\odot$  are expected to be, one could substantially underestimate  $n$  and  $R_e$  for the typical  $S/N$  ( $\leq 50$ – $100$ ) of their sample. They believe that this reflects the impact of  $(1+z)^4$  surface brightness

TABLE 7  
MULTI-COMPONENT GALAXIES WITH OBSERVED BACKGROUND – CORRECTED FOR MODELS WITH  $n > 6$   
 $\Delta R_e/R_e$

(1)	Subsample (2)	$S/N$							
		5 – 10 (3)	10 – 20 (4)	20 – 35 (5)	35 – 70 (6)	70 – 135 (7)	135 – 265 (8)	265 – 515 (9)	515 – 1000 (10)
$R_e$	1 – 5	$-0.46^{+1.12}_{-0.54}$	$-0.36^{+2.07}_{-0.49}$	$-0.23^{+1.03}_{-0.39}$	$-0.14^{+0.56}_{-0.40}$	$-0.05^{+0.28}_{-0.37}$	$0.02^{+0.19}_{-0.49}$	$0.08^{+0.09}_{-0.53}$	$0.05^{+0.08}_{-0.44}$
	5 – 20	$-0.49^{+1.49}_{-0.42}$	$-0.30^{+1.81}_{-0.38}$	$-0.04^{+1.05}_{-0.44}$	$0.13^{+0.66}_{-0.43}$	$0.13^{+0.49}_{-0.28}$	$0.18^{+0.33}_{-0.25}$	$0.23^{+0.27}_{-0.27}$	$0.21^{+0.26}_{-0.23}$
	20 – 70	$-0.43^{+1.30}_{-0.37}$	$-0.28^{+0.99}_{-0.41}$	$-0.00^{+0.72}_{-0.51}$	$-0.02^{+0.50}_{-0.42}$	$0.03^{+0.41}_{-0.45}$	$0.06^{+0.49}_{-0.35}$	$0.10^{+0.45}_{-0.42}$	$0.16^{+0.41}_{-0.35}$
$\langle\mu\rangle$	16 – 21	...	...	...	$-0.51^{+0.17}_{-0.29}$	$-0.47^{+0.04}_{-0.12}$	$-0.08^{+0.26}_{-0.40}$	$0.13^{+0.23}_{-0.21}$	$0.15^{+0.31}_{-0.20}$
	21 – 24	$-0.86^{+0.77}_{-0.14}$	$-0.38^{+1.45}_{-0.45}$	$-0.05^{+0.96}_{-0.39}$	$0.03^{+0.58}_{-0.39}$	$0.07^{+0.47}_{-0.33}$	$0.12^{+0.39}_{-0.25}$	$0.22^{+0.34}_{-0.35}$	$0.19^{+0.30}_{-0.38}$
	24 – 27	$-0.47^{+1.46}_{-0.41}$	$-0.28^{+1.38}_{-0.42}$	$-0.05^{+0.92}_{-0.46}$	$0.07^{+0.61}_{-0.47}$	$0.04^{+0.44}_{-0.27}$	...	...	...
$n$	0 – 2.5	$-0.60^{+0.57}_{-0.24}$	$-0.57^{+0.31}_{-0.22}$	$-0.47^{+0.28}_{-0.22}$	$-0.45^{+0.25}_{-0.18}$	$-0.27^{+0.24}_{-0.29}$	$-0.15^{+0.09}_{-0.31}$	$-0.24^{+0.16}_{-0.34}$	$-0.19^{+0.09}_{-0.29}$
	2.5 – 6	$-0.25^{+1.92}_{-0.58}$	$-0.19^{+1.18}_{-0.42}$	$-0.09^{+0.62}_{-0.37}$	$0.02^{+0.47}_{-0.34}$	$0.05^{+0.39}_{-0.29}$	$0.11^{+0.35}_{-0.21}$	$0.17^{+0.28}_{-0.20}$	$0.15^{+0.25}_{-0.16}$
	6	$-0.21^{+2.24}_{-0.79}$	$0.22^{+1.82}_{-0.79}$	$0.56^{+1.23}_{-0.75}$	$0.37^{+0.78}_{-0.45}$	$0.26^{+0.54}_{-0.42}$	$0.27^{+0.40}_{-0.37}$	$0.30^{+0.33}_{-0.43}$	$0.33^{+0.30}_{-0.36}$

NOTE. — Col. (1) Parameter. Col. (2) The subsamples selected based on different ranges of effective radius ( $R_e$ ), surface brightness ( $\langle\mu\rangle$ ), and Sérsic index ( $n$ ). Col. (3)-(10) Equally (in logarithmic space) spaced intervals of  $S/N$ . See text in Section 3.2 regarding the offsets in different Sérsic index.

TABLE 8  
MULTI-COMPONENT GALAXIES WITH OBSERVED BACKGROUND – CORRECTED FOR MODELS WITH  $n > 6$   
 $\Delta m_H$

(1)	Subsample (2)	$S/N$							
		5 – 10 (3)	10 – 20 (4)	20 – 35 (5)	35 – 70 (6)	70 – 135 (7)	135 – 265 (8)	265 – 515 (9)	515 – 1000 (10)
$R_e$	1 – 5	$0.04^{+0.78}_{-1.32}$	$0.07^{+0.44}_{-0.96}$	$0.14^{+0.37}_{-0.53}$	$0.05^{+0.37}_{-0.26}$	$0.07^{+0.22}_{-0.22}$	$-0.01^{+0.40}_{-0.11}$	$-0.05^{+0.29}_{-0.06}$	$-0.03^{+0.25}_{-0.06}$
	5 – 20	$0.11^{+0.76}_{-1.09}$	$0.10^{+0.56}_{-0.73}$	$-0.01^{+0.38}_{-0.45}$	$-0.08^{+0.30}_{-0.27}$	$-0.07^{+0.17}_{-0.20}$	$-0.10^{+0.12}_{-0.14}$	$-0.12^{+0.12}_{-0.11}$	$-0.11^{+0.12}_{-0.11}$
	20 – 70	$0.41^{+0.68}_{-0.94}$	$0.23^{+0.55}_{-0.55}$	$0.03^{+0.45}_{-0.35}$	$0.04^{+0.36}_{-0.24}$	$0.01^{+0.37}_{-0.20}$	$-0.02^{+0.23}_{-0.24}$	$-0.04^{+0.27}_{-0.21}$	$-0.09^{+0.23}_{-0.17}$
$\langle\mu\rangle$	16 – 21	...	...	...	$0.39^{+0.19}_{-0.22}$	$0.40^{+0.14}_{-0.15}$	$0.04^{+0.38}_{-0.14}$	$-0.08^{+0.14}_{-0.10}$	$-0.09^{+0.13}_{-0.12}$
	21 – 24	$0.51^{+5.43}_{-0.21}$	$0.28^{+0.38}_{-0.89}$	$-0.00^{+0.38}_{-0.38}$	$-0.02^{+0.29}_{-0.28}$	$-0.03^{+0.23}_{-0.21}$	$-0.07^{+0.15}_{-0.16}$	$-0.11^{+0.18}_{-0.14}$	$-0.10^{+0.26}_{-0.14}$
	24 – 27	$0.16^{+0.76}_{-1.03}$	$0.13^{+0.53}_{-0.69}$	$0.02^{+0.41}_{-0.44}$	$-0.04^{+0.37}_{-0.28}$	$-0.02^{+0.20}_{-0.24}$	...	...	...
$n$	0 – 2.5	$0.34^{+0.62}_{-0.83}$	$0.48^{+0.42}_{-0.43}$	$0.40^{+0.31}_{-0.31}$	$0.41^{+0.22}_{-0.30}$	$0.24^{+0.33}_{-0.23}$	$0.16^{+0.30}_{-0.14}$	$0.27^{+0.33}_{-0.26}$	$0.22^{+0.36}_{-0.17}$
	2.5 – 6	$-0.02^{+0.79}_{-1.03}$	$0.06^{+0.43}_{-0.60}$	$0.03^{+0.36}_{-0.30}$	$-0.01^{+0.25}_{-0.24}$	$-0.03^{+0.20}_{-0.18}$	$-0.07^{+0.13}_{-0.15}$	$-0.10^{+0.12}_{-0.12}$	$-0.09^{+0.10}_{-0.11}$
	6	$0.02^{+1.00}_{-1.12}$	$-0.16^{+0.54}_{-0.65}$	$-0.24^{+0.41}_{-0.42}$	$-0.17^{+0.22}_{-0.27}$	$-0.12^{+0.20}_{-0.23}$	$-0.11^{+0.23}_{-0.18}$	$-0.13^{+0.18}_{-0.14}$	$-0.15^{+0.16}_{-0.12}$

NOTE. — Col. (1) Parameter. Col. (2) The subsamples selected based on different ranges of effective radius ( $R_e$ ), surface brightness ( $\langle\mu\rangle$ ), and Sérsic index ( $n$ ). Col. (3)-(10) Equally (in logarithmic space) spaced intervals of  $S/N$ . See text in Section 3.2 regarding the offsets in different Sérsic index.

dimming of high- $n$  halos. We do not observe any systematic error in size determination of single-component galaxies with  $S/N > 50$  ( $m_H < 24$  for typical  $z = 2$  galaxies).

The sizes of single-component galaxies can be underestimated if the dimension of the image is small relative to the galaxy size. This makes the sky determination less reliable. Size measurements can also be compromised if extended neighbors are not properly masked. However, when these factors are taken into account, we, in agreement with several previous studies (Häussler et al. 2007; van Dokkum & Brammer 2010; Williams et al.

2010; Buitrago et al. 2013), believe that the parameters of a single-component galaxy can be retrieved robustly over a wide range of  $S/N$ .

Meert et al. (2013) show that single-component Sérsic fits of two-component (bulge + disk) simulated galaxies, based on a spectroscopically selected sample from the Sloan Digital Sky Survey (SDSS; York et al. 2000; Stoughton et al. 2002), leads to an overestimation of the size (their Figure 8). It holds even for their model galaxies without any noise (their Figure 8a).

Mosleh et al. (2013) also simulate two-component galaxies based on an SDSS sample. They examine the ro-

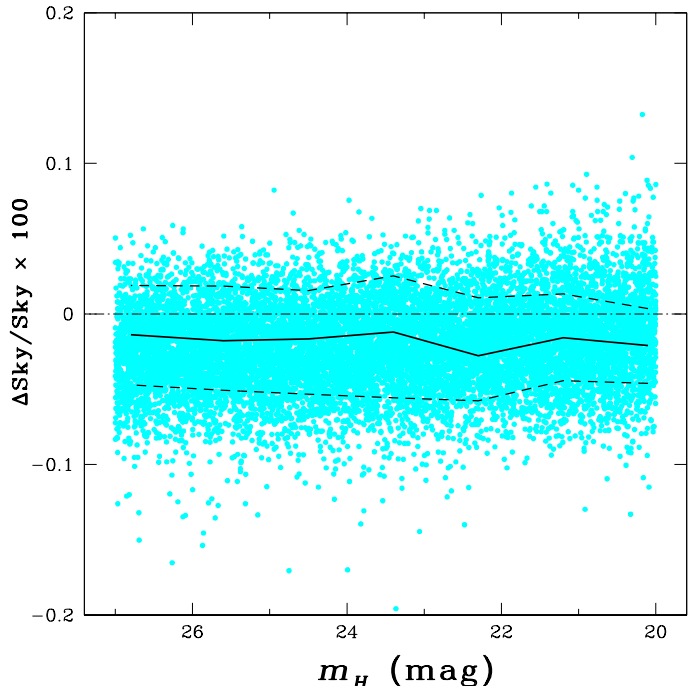


FIG. 9.— Examining the accuracy of GALFIT sky determination. Black solid and dashed lines in the scatter plots indicate the median and  $1\sigma$  uncertainties of different measurements. The dot-dash line shows zero offset. For our simulations, GALFIT measures the sky value with accuracy better than a few 0.01%.

bustness of size measurements based on non-parametric fitting as well as single- and two-component Sérsic fits. For their modeled galaxies at  $z = 0$ , they find that their non-parametric and two-component Sérsic fits provide the most robust  $R_e$  measurements, while those based on single-component Sérsic fitting often overestimate the size, especially for massive red/early-type galaxies (their Figure 3). Redshifting their galaxies to  $z = 1$ , they find that the single-Sérsic fitting of two-component galaxies yield reliable size measurements, likely due to the smaller structures being washed out at high redshifts (their Figure 7).

The two-component simulation results of Meert et al. (2013) and Mosleh et al. (2013) qualitatively agree with our results for multi-component galaxies (Figures 6–8). However, Mosleh et al. (2013) finds smaller systematic offsets which could be caused by a considerable fraction of disk-dominated galaxies in their sample. Note that our simulations are limited to early-type galaxies.

## 5. IMPLICATIONS FOR RED NUGGETS

One of the main goals of this study is to examine whether or not there is a bias in measuring properties of the red, massive compact galaxies at  $z = 2$ . These so-called red nuggets are found to have a median stellar mass  $M_* \approx 10^{11} M_\odot$  and a median effective radius  $R_e \leq 2.0$  kpc (van Dokkum et al. 2008; van der Wel et al. 2011). Red nuggets are found to have  $H_{160} < 24$ , with a few having  $H_{160} = 23$ –24 (Szomoru et al. 2012). They are seen at high  $S/N$  ( $\geq 100$ , for one *HST* orbit).

Assuming that the light distribution of these compact galaxies resembles a Sérsic profile, Figures 4 and 5 and Tables 1 and 2 show that there is no systematic bias

in the size, total luminosity, Sérsic index, and ellipticity of these galaxies. The uncertainties for measuring  $R_e$ ,  $m_H$ ,  $n$ , and  $e$  are less than 20%, 0.2, 0.2, and 10%, respectively, for galaxies with  $23 < H_{160} < 24$ . The uncertainties are much smaller still for brighter galaxies. Thus, provided that red nuggets have Sérsic profiles, single-component fits of these galaxies are very robust.

On the other hand, the observed compact galaxies may have an envelope with low surface brightness that is difficult to recover in detail. If the envelope does not follow a perfect Sérsic function or the overall galaxy profile changes from small to large radii, it is commonly speculated that the envelope flux and hence the true effective radius may be underestimated (e.g., Hopkins et al. 2009). We tackle this problem by taking advantage of H13’s morphological study of  $\sim 100$  nearby massive elliptical galaxies with median mass  $M_\odot = 1.3 \times 10^{11} M_\odot$ . We directly test the null hypothesis that the red nuggets have undergone no morphological evolution and only passive fading since  $z \approx 2$ .

Figure 7 and Table 5 and 6 show the best-fit results of the single Sérsic fits of the multi-component galaxies. With adequate  $S/N$  ( $\geq 100$ , comparable to red nuggets studied in CANDELS), the presence of extended envelopes in ellipticals actually leads to a slight *overestimation* of their sizes rather than an underestimation. This holds over a wide range of sizes (i.e.,  $1 < R_e < 70$  pixels). At much lower  $S/N$  values (e.g.,  $\leq 50$ ), the trend reverses but the effect is very modest. Conclusion: if red nuggets have structures similar to local massive elliptical galaxies, single-component Sérsic fits of these galaxies do not underestimate their sizes. The median size of models of nearby ellipticals rescaled to mimic galaxies at  $z = 2$  is  $R_e \approx 0.5$ , or 4.2 kpc. Considering that the median size of red nuggets is  $R_e \approx 1$  kpc, their sizes have increased by a factor of  $\sim 4$  if they are the progenitors of local massive ellipticals.

Table 6 shows that the total luminosity of the multi-component galaxies can be overestimated as a result of overestimating the size. This leads to systematically higher masses for these compact galaxies. However, this effect is less than 10%, which, compared to other uncertainties involved in mass measurement, is insignificant. Hence, the total luminosity of multi-component galaxies can be measured reliably via single-component Sérsic fitting.

## 6. SUMMARY

Recent observations of galaxies at  $z = 2$  with *HST*/WFC3 have revealed a population of red, compact ( $R_e \approx 1$ –2 kpc), and massive ( $M_* \approx 10^{11} M_\odot$ ) galaxies, the so-called red nuggets. We determine the possible biases and uncertainties in the determination of basic structural parameters of these galaxies, with special emphasis on their sizes.

For this purpose, we perform two sets of simulations: one based on idealized single-component models and the other on the observed properties of local massive ellipticals. For the local analogs, we generate galaxies with multiple components (generally three) with the properties obtained from Huang et al. (2013a), which in turn are rescaled so as to compare with galaxies observed at  $z = 2.0$ . We examine the effects of background noise, the accuracy of the PSF model, and the model fitting

method. We analyze the artificial images in a manner similar to that popularly employed in the literature, namely through two-dimensional **GALFIT** modeling of the light distribution using a single-component Sérsic profile and compare the retrieved size, luminosity, and other structural parameters with input values.

We find that:

- If the sky estimation has been done robustly and the PSF model is relatively accurate, **GALFIT** can retrieve the properties of single-component galaxies at a wide range of magnitudes without introducing any systematic error at high accuracy (Figures 4 and 5 and Tables 1 and 2).
- Modeling multi-component galaxies with a single Sérsic component under realistic conditions does not bias the sizes too low; in fact, sizes tend to be slightly overestimated. This makes the size evolution problem, if anything, even more dramatic. The apparent compactness of red nuggets is real; it is not the result of missing faint, outer light (Figures 6–8 and Table 5).
- Models with Sérsic indices larger than 6 have larger uncertainties and can cause significant systematic

errors. Refitting these galaxies by fixing the Sérsic index to  $n = 6$  provides more reliable results (compare Figure 6 with Figure 7).

- We confirm that massive, compact quiescent galaxies at  $z = 2$  are a factor of  $\sim 4$  smaller than their local counterparts.

We thank the anonymous referee for a careful reading of the paper and for making detailed suggestions that improved the understanding of this work. RD has been funded by a graduate student fellowship awarded by Carnegie Observatories; we are grateful to director Wendy Freedman for her support. LCH acknowledges support from the Kavli Foundation, Peking University, the Chinese Academy of Sciences, and the Carnegie Institution for Science. RD would like to thank Prof. G. Canalizo, and Heather L. Worthington for providing long-term support and Minjin Kim, Moein Mosleh, Louis E. Abramson, and Shoubaneh Hemmati for useful discussions. This work was partly funded by NASA grants HST-AR-12818 and HST-GO-12903 from the Space Telescope Science Institute (operated by AURA, Inc., under NASA contract NAS5-26555).

## APPENDIX

### *The Effects of PSF Accuracy*

An important factor in obtaining a best-fit model is the accuracy of the PSF. The effective radii of high- $z$  galaxies are comparable to and in some cases smaller than the FWHM of the PSF. Therefore, one may expect a considerable offset in **GALFIT** measurements when an inaccurate PSF is used (i.e., a PSF with different FWHM and/or different structure than the PSF used for generating the galaxy model). To study this effect, we first generate about 3,000 single-component galaxies with  $0.8 < R_e < 4$  pixels, which are convolved with the CANDELS COSMOS  $H_{160}$  PSF. Then, we fit each galaxy with the  $H_{160}$  PSF from several major *HST* surveys (CANDELS, COSMOS, UDS, GOODS-S, and EGS) and compare the results. The PSFs use a combination of a hybrid TinyTim (Krist et al. 2011) PSF and a stacked empirical stellar PSF. The motivation for constructing such a hybrid PSF is that the TinyTim PSFs are better in the core region (where centroiding issues tend to broaden empirical PSFs), while empirical PSFs fit real stars better in the wings (van der Wel et al. 2012). Figure A1 shows that these PSF models have different core sizes (i.e. FWHM) and outer halo wing compared to the COSMOS PSF. The PSF in COSMOS has an intermediate core size.

Figure A2 shows that the exact choice of PSF, within the range of realistic PSF models we explored, is not important for the purposes of galaxy size determination. The systematic offsets are less than 2% and the uncertainties are less than 4%. The offsets are smaller for best-fit models with lower Sérsic indices.



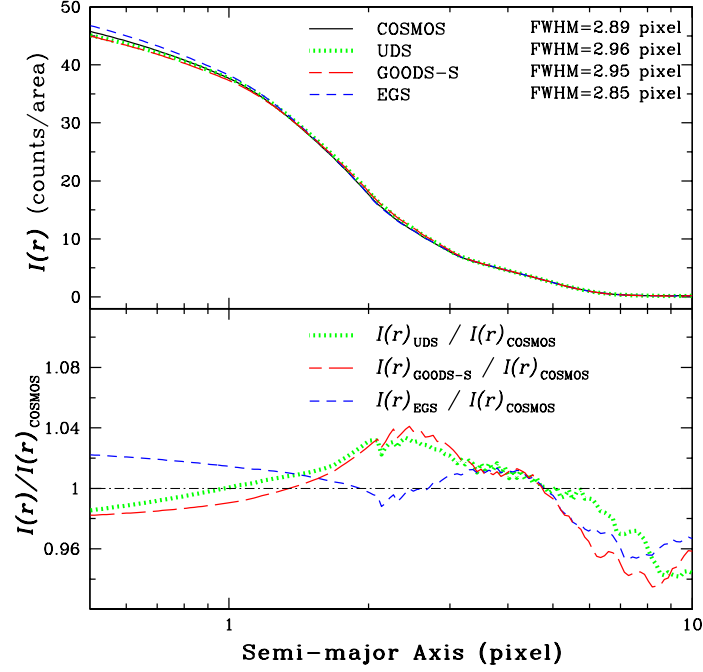


FIG. A1.— CANDELS  $H_{160}$  PSF models for different fields. The PSF models are constructed by using TinyTim for the central and outer parts of the PSF and bright stars for intermediate distances from the center. The top panel shows the one-dimensional light profile of different PSF models. The bottom panel shows the differences between the structure of different PSFs. The dot-dash line marks the case when there is no offset.

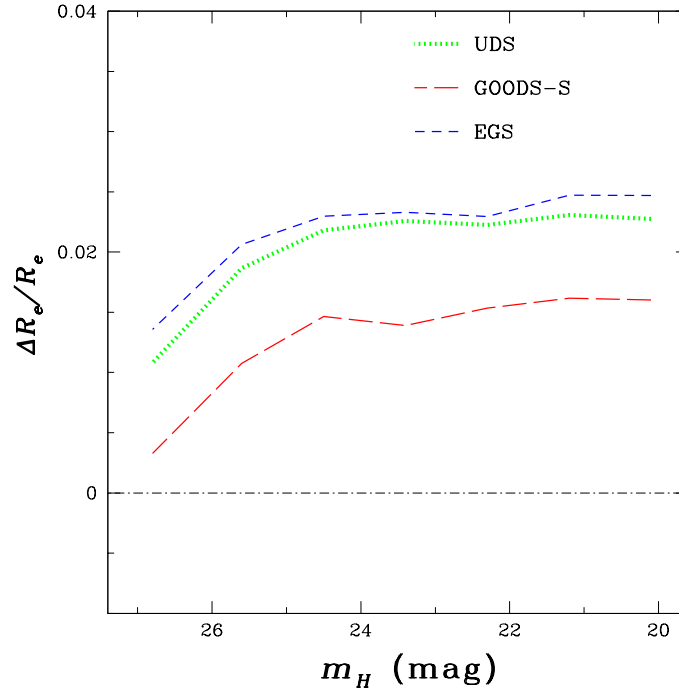


FIG. A2.— Fitting galaxies with PSFs with different structures and/or FWHMs than the intrinsic PSF of the image leads to biases that are insignificant. The systematic offsets for  $R_e$  are less than 2%, and the uncertainties are less than 4%. The dot-dash line shows zero offset.

## REFERENCES

- Abramson, L. E., Dressler, A., Gladders, M. D., et al. 2013, *ApJ*, 777, 124
- Barden, M., Rix, H.-W., Somerville, R. S., et al. 2005, *ApJ*, 635, 959
- Barden, M., Häußler, B., Peng, C. Y., McIntosh, D. H., & Guo, Y. 2012, *MNRAS*, 422, 449
- Bell, E. F., Wolf, C., Meisenheimer, K., et al. 2004, *ApJ*, 608, 752
- Bernardi, M., Shankar, F., Hyde, J. B., et al. 2010, *MNRAS*, 404, 2087
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
- Blanton, M. R., Hogg, D. W., Bahcall, N. A., et al. 2003, *ApJ*, 592, 819
- Bruce, V. A., Dunlop, J. S., Cirasuolo, M., et al. 2012, *MNRAS*, 427, 1666
- Buitrago, F., Trujillo, I., Conselice, C. J., et al. 2008, *ApJ*, 687, L61
- Buitrago, F., Trujillo, I., Conselice, C. J., Häussler, B. 2013, *MNRAS*, 428, 1460
- Cassata, P., Giavalisco, M., Guo, Y., et al. 2010, *ApJ*, 714, L79
- Cassata, P., Giavalisco, M., Guo, Y., et al. 2011, *ApJ*, 743, 96
- Cimatti, A., Cassata, P., Pozzetti, L., et al. 2008, *A&A*, 482, 21
- Ciotti, L. 1991, *A&A*, 249, 99
- Conselice, C. J. 1997, *PASP*, 109, 1251
- Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, *ApJ*, 626, 680
- Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, *ApJ*, 695, 101
- de Vaucouleurs, G. 1948, *Annales d'Astrophysique*, 11, 247
- Erwin, P., Pohlen, M., & Beckman, J. E. 2008, *AJ*, 135, 20
- Fisher, D. B., & Drory, N. 2008, *AJ*, 136, 773
- Franx, M., Labbé, I., Rudnick, G., et al. 2003, *ApJ*, 587, L79
- Franx, M., van Dokkum, P. G., Schreiber, N. M. F., et al. 2008, *ApJ*, 688, 770
- Freeman, K. C. 1970, *ApJ*, 160, 811
- Fruchter, A., Sosey, M., Hack, W., et al. 2009, *HST MultiDrizzle Handbook*
- Graham, A. W., Driver, S. P., Petrosian, V., et al. 2005, *AJ*, 130, 1535
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 35
- Häussler, B., McIntosh, D. H., Barden, M., et al. 2007, *ApJS*, 172, 615
- Ho, L. C., Li, Z.-Y., Barth, A. J., Seigar, M. S., & Peng, C. Y. 2011, *ApJS*, 197, 21
- Hogg, D. W., Blanton, M. R., Brinchmann, J., et al. 2004, *ApJ*, 601, L29
- Hopkins, P. F., Bundy, K., Murray, N., et al. 2009, *MNRAS*, 398, 898
- Huang, S., Ho, L. C., Peng, C. Y., Li, Z.-Y., & Barth, A. J. 2013a, *ApJ*, 766, 47
- Huang, S., Ho, L. C., Peng, C. Y., Li, Z.-Y., & Barth, A. J. 2013b, *ApJ*, 768, L28
- Ichikawa, T., Kajisawa, M., & Akhlaghi, M. 2012, *MNRAS*, 422, 1014
- Jedrzejewski, R. I. 1987, *MNRAS*, 226, 747
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36
- Kriek, M., van Dokkum, P. G., Franx, M., et al. 2006, *ApJ*, 649, L71
- Krist, J. E., Hook, R. N., & Stoehr, F. 2011, *SPIE*, 8127, 16
- Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, 379, 1599
- Lotz, J. M., Primack, J., & Madau, P. 2004, *AJ*, 128, 163
- MacArthur, L. A., Courteau, S., & Holtzman, J. A. 2003, *ApJ*, 582, 689
- Mancini, C., Daddi, E., Renzini, A., et al. 2010, *MNRAS*, 401, 933
- McIntosh, D. H., Bell, E. F., Rix, H.-W., et al. 2005, *ApJ*, 632, 191
- Meert, A., Vikram, V., & Bernardi, M. 2013, *MNRAS*, 433, 1344
- Mosleh, M., Williams, R. J., & Franx, M. 2013, *ApJ*, 777, 117
- Muzzin, A., van Dokkum, P., Franx, M., et al. 2009, *ApJ*, 706, L188
- Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, *ApJ*, 746, 162
- Oser, L., Naab, T., Ostriker, J. P., & Johansson, P. H. 2012, *ApJ*, 744, 63
- Papovich, C., Bassett, R., Lotz, J. M., et al. 2012, *ApJ*, 750, 93
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, *AJ*, 124, 266
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, *AJ*, 139, 2097
- Poggianti, B. M., Calvi, R., Bindoni, D., et al. 2013, *ApJ*, 762, 77
- Ravindranath, S., Ferguson, H. C., Conselice, C., et al. 2004, *ApJ*, 604, L9
- Saracco, P., Longhetti, M., & Gargiulo, A. 2010, *MNRAS*, 408, L21
- Shen, S., Mo, H. J., White, S. D. M., et al. 2003, *MNRAS*, 343, 978
- Simard, L., & Pritchet, C. J. 1998, *ApJ*, 505, 96
- Simard, L., Willmer, C. N. A., Vogt, N. P., et al. 2002, *ApJS*, 142, 1
- Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, *ApJS*, 196, 11
- Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, *AJ*, 123, 485
- Szomoru, D., Franx, M., van Dokkum, P. G., et al. 2010, *ApJ*, 714, L244
- Szomoru, D., Franx, M., & van Dokkum, P. G. 2012, *ApJ*, 749, 121
- Taylor, E. N., Franx, M., Glazebrook, K., et al. 2010, *ApJ*, 720, 723
- Tody, D. 1986, *SPIE*, 627, 733
- Tody, D. 1993, *Astronomical Society of the Pacific Conference Series*, 52, 173
- Toft, S., van Dokkum, P., Franx, M., et al. 2007, *ApJ*, 671, 285
- Trujillo, I., Conselice, C. J., Bundy, K., et al. 2007, *MNRAS*, 382, 109
- Trujillo, I., Cenarro, A. J., de Lorenzo-Cáceres, A., et al. 2009, *ApJ*, 692, L118
- Valentinuzzi, T., Fritz, J., Poggianti, B. M., et al. 2010, *ApJ*, 712, 226
- van der Wel, A., Holden, B. P., Zirm, A. W., et al. 2008, *ApJ*, 688, 48
- van der Wel, A., Rix, H.-W., Wuyts, S., et al. 2011, *ApJ*, 730, 38
- van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, *ApJS*, 203, 24
- van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, *ApJ*, 677, L5
- van Dokkum, P. G., & Brammer, G. 2010, *ApJ*, 718, L73
- Williams, R. J., Quadri, R. F., Franx, M., et al. 2010, *ApJ*, 713, 738
- Wuyts, S., Cox, T. J., Hayward, C. C., et al. 2010, *ApJ*, 722, 1666
- Yoon, I., Weinberg, M. D., & Katz, N. 2011, *MNRAS*, 414, 1625
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579